

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:50 ; Search time 60 Seconds
(without alignments)
1186.700 Million cell updates/sec

Title: US-09-857-826b-17
Perfect score: 252
Sequence: 1 MAELFVQIIIVVMVWV.....PLESAIWSKEDKQKHPL 252

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 29Jan04.*
1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001s.*
5: geneseq2002s.*
6: geneseq2003as.*
7: geneseq2003bs.*
8: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252	100.0	252	3 AAB01388	Aab01388 Neuron-as
2	252	100.0	252	3 AAB18449	Aab18449 A human T
3	252	100.0	252	5 AAG61805	Abg61805 Prostate
4	252	100.0	252	6 ABP97219	Abp97219 Tumour-as
5	252	100.0	252	6 ABP97218	Abp97218 Tumour-as
6	252	100.0	252	7 ADB75589	Adb75589 Prostate
7	252	100.0	252	7 ADC37325	Adc37325 Nuclear f
8	252	100.0	252	7 AAE39589	Aae39589 Human and
9	252	100.0	252	6 ABE75660	Abp75660 Human sec
10	250	99.2	287	5 AAU78231	Aau78231 Human Mec
11	250	99.2	287	6 ABP97234	Abp97234 Tumour-as
12	250	99.2	287	7 ADC37327	Adc37327 Nuclear f
13	249	98.8	249	7 AAE39591	Aae39591 Human and
14	249	98.8	252	3 AAB18461	Aab18461 A human T
15	237	94.0	232	4 AAM18712	Aam18712 Human pol
16	207	82.1	207	4 AAM40498	Aam40498 Human pol
17	204	81.0	252	3 AAB18462	Aab18462 A human T
18	198	78.6	252	3 AAB18463	Aab18463 A human T
19	181	71.8	241	5 ABG77052	Abg77052 Prostate
20	106	42.1	270	2 AAY74136	Aay74136 Human pro
21	51	20.2	217	3 AAB18466	Aab18466 A murine
22	51	20.2	217	3 AAB18464	Aab18464 A murine
23	51	20.2	217	3 AAB18465	Aab18465 A murine
24	51	20.2	217	3 AAB18450	Aab18450 A murine
25	51	20.2	274	5 AAU78236	Aau78236 Mouse Mec

26	35	13.9	261	4	ABG20316	Abg20316 Novel hum
27	35	13.9	269	3	AAV73422	Aay73422 Human sec
28	35	13.9	269	6	ADA54268	Ada54268 Human pro
29	35	13.9	288	2	AAW60674	Aaw60674 Manic-dep
30	35	13.9	288	7	ADC37453	Adc37453 Nucleic f
31	35	13.9	306	2	AAW60673	Aaw60673 Manic-dep
32	35	13.9	306	7	ADE59611	Ades9611 Human pro
33	35	13.9	316	4	ABG13118	Abg13118 Novel hum
34	35	13.9	334	4	ABG13117	Abg13117 Novel hum
35	35	13.9	475	4	ABG28310	Abg28310 Novel hum
36	35	13.9	485	4	ABG13126	Abg13126 Novel hum
37	20	7.9	204	4	ABG13125	Abg13125 Novel hum
38	20	7.9	235	4	ABG13123	Abg13123 Novel hum
39	14	5.6	14	4	AAW00597	Aam00597 Human pro
40	9	3.6	19	4	ABG20315	Abg20315 Novel hum
41	8	3.2	308	4	ABE71424	Abb71424 Drosophil
42	8	3.2	308	4	ABB66489	Abb66489 Drosophil
43	8	3.2	308	4	ABB66488	Abb66488 Drosophil
44	7	2.8	14	4	AAW00598	Aam00598 Human pro
45	7	2.8	32	4	AAW15296	Aam15296 Peptide #

ALIGNMENTS

RESULT 1
ID AAB01388 standard; protein; 252 AA.

XX AAB01388;

DT 20-OCT-2000 (first entry)

DE Neuron-associated protein.

KW Neuron associated protein; NEURAP; neurological disorder; epilepsy;
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
KW Parkinson's disease; demyelinating disease; meningitis; prion disease;
KW kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;
KW muscular dystrophy; central nervous system; CNS;
KW peripheral nervous system; PNS; myopathy; schizophrenia;
KW actinic keratosis; arteriosclerosis; atherosclerosis; MCTD;
KW cirrhosis; hepatitis; mixed connective tissue disease; cancer;
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KW AIDS; Addison's disease; adult respiratory distress syndrome; allergy;
KW ankylosing spondylitis; amyloidosis; anaemia; asthma;
KW Werner syndrome; trauma; human.

XX Homo sapiens.

Key Location/Qualifiers

Peptide 1..33

Region 9..27

FT /label= Signal peptide

FT /label= Transmembrane region

FT Modified-site 33

FT /note= "Potential phosphorylation site"

FT Modified-site 65..68

FT /note= "Glycosaminoglycan attachment site"

FT Modified-site 81

FT /note= "Potential phosphorylation site"

FT Modified-site 116

FT /note= "Potential phosphorylation site"

FT Modified-site 118

FT /note= "Potential phosphorylation site"

FT Modified-site 147

FT /note= "Potential phosphorylation site"

FT Modified-site 153

FT /note= "Potential glycosylation site"

FT Modified-site 155

FT /note= "Potential phosphorylation site"

CC cells that appear in the bone marrow, and leukocytes, treat bone marrow,
CC blood and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to treat
CC disorders associated with the ovaries, cerebral oedema, hydrocephalus,
CC brain herniations, iatrogenic disease, inflammations, bacterial and viral
CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
CC disease, multiple sclerosis, brain cancers, hydrocephalus and
CC encephalitis, and treat hepatic disorders

XX Sequence 252 AA;

Query Match 100.0%; Score 252; DB 3; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.8e-233;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
Db 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRPQTPYQLQHEIDLPTTISLSDG 120
Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRPQTPYQLQHEIDLPTTISLSDG 120
QY 121 BEPPYQGPCTQLQDRDPEQQLNEESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Db 121 BEPPYQGPCTQLQDRDPEQQLNEESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
QY 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
QY 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 3
ABG61805
ID ABG61805 standard; protein; 252 AA.

XX AC ABG61805;

XX 15-AUG-2002 (first entry)

XX Prostate cancer-associated protein #6.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

XX 08-DEC-2000; 2000US-00733288.

XX 08-DEC-2000; 2000US-00733742.

XX 24-JAN-2001; 2001US-0263957P.

XX 16-MAR-2001; 2001US-0276791P.

XX 16-MAR-2001; 2001US-0276888P.

XX 06-APR-2001; 2001US-0281922P.

XX 24-APR-2001; 2001US-0286214P.

XX 30-APR-2001; 2001US-00847046.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
PI WPI; 2002-471335/50.
XX N-PSDB; ABK92120.
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
XX useful for diagnosing prostate cancer (PC) or screening modulators of PC,
XX by determining if prostate cancer-associated genes are expressed in a
XX prostate tissue.
XX Claim 27; Page 305; 436pp; English.
XX The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
XX associated polynucleotides (designated PC genes) that selectively
XX hybridise to a sequence that is at least 80% identical to them. The
XX prostate cancer-associated polynucleotide sequences are differentially
XX expressed in prostate tumour tissue or in prostate cancer and are derived
XX from the tissues of various organisms such as humans or other mammals
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX associated genes are useful for diagnosing or treating prostate cancer,
XX as well as for identifying modulators of prostate cancer or agents that
XX inhibit prostate cancer. The nucleic acid sequences are particularly
XX useful in gene therapy, as a vaccine or in antisense applications.
XX ABG61800-ABG61944 represent prostate cancer-associated proteins
XX Sequence 252 AA;
SQ

Query Match 100.0%; Score 252; DB 5; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.8e-233;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
Db 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRPQTPYQLQHEIDLPTTISLSDG 120
Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRPQTPYQLQHEIDLPTTISLSDG 120
QY 121 BEPPYQGPCTQLQDRDPEQQLNEESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Db 121 BEPPYQGPCTQLQDRDPEQQLNEESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
QY 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
QY 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 4

ABP97219

ID ABP97219 standard; protein; 252 AA.

XX AC ABP97219;

XX 01-JUL-2003 (first entry)

XX Tumour-associated antigenic target protein TAT375 SEQ ID NO:101.

XX Human; tumour-associated antigenic target; TAT; tumour; diagnosis;

XX Homo sapiens.

XX WO2003024392-A2.

XX

```

PD 27-MAR-2003.
XX
XX
XX 11-SEP-2002; 2002WO-US028859.
XX
XX 18-SEP-2001; 2001US-0323268P.
XX 19-OCT-2001; 2001US-0339227P.
XX 07-NOV-2001; 2001US-0336827P.
XX 20-NOV-2001; 2001US-0331906P.
XX 02-JAN-2002; 2002US-0345444P.
XX 03-APR-2002; 2002US-0369724P.
XX 19-AUG-2002; 2002US-0404809P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
XX Williams PM, Wu TD, Zhang Z;
XX
XX WPI; 2003-354551/33.
XX N-PSDB; ACC49537.
XX
XX New antibodies against tumor-associated antigenic target polypeptide,
XX useful for treating or diagnosing tumors or cancers in mammals, e.g.
XX prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
XX carcinomas.
XX
XX Claim 2; Fig 101; 285pp; English.
XX
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target
XX (TAT) proteins given in ABP97175 to ABP97234. The present invention
XX describes an isolated antibody that binds to a polypeptide having at
XX least 80 % sequence identity to any of the 60 150-800 residue amino acid
XX sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
XX its associated signal peptide, encoded by any of the 60 2000-3000 base
XX pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
XX cytostatic activity. The antibody can be used for treating or diagnosing
XX tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
XX cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
XX cell carcinomas, or thyroid cancer
XX
XX Sequence 252 AA;
XX
XX Query Match 100.0%; Score 252; DB 6; Length 252;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-233;
XX Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPFISHSQRRREDALSSGCLMPS 60
XX Db 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPFISHSQRRREDALSSGCLMPS 60
XX
XX QY 61 ESTVSGNGIPEPQVYAPPRPTDLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
XX Db 61 ESTVSGNGIPEPQVYAPPRPTDLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
XX
XX QY 121 BEPPPYQGCTQLQRPDPEQQLNRESVRAPNRTTFDSLDMSARLGCGPCPPSSNSGIS 180
XX Db 121 BEPPPYQGCTQLQRPDPEQQLNRESVRAPNRTTFDSLDMSARLGCGPCPPSSNSGIS 180
XX
XX QY 181 ATCYGSGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
XX Db 181 ATCYGSGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
XX
XX QY 241 SKEKDQKQHPL 252
XX Db 241 SKEKDQKQHPL 252
XX
XX RESULT 5
XX ID ABP97218
XX ID ABP97218 standard; protein; 252 AA.
XX AC ABP97218;
XX
XX DT 01-JUL-2003 (first entry)

```

```

XX
XX DE Tumour-associated antigenic target protein TAT180 SEQ ID NO:100.
XX
XX KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
XX KW Cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO2003024392-A2.
XX
XX PD 27-MAR-2003.
XX
XX PF 11-SEP-2002; 2002WO-US028859.
XX
XX PR 18-SEP-2001; 2001US-0323268P.
XX PR 19-OCT-2001; 2001US-0339227P.
XX PR 07-NOV-2001; 2001US-0336827P.
XX PR 20-NOV-2001; 2001US-0331906P.
XX PR 02-JAN-2002; 2002US-0345444P.
XX PR 03-APR-2002; 2002US-0369724P.
XX PR 19-AUG-2002; 2002US-0404809P.
XX
XX PA (GETH ) GENENTECH INC.
XX
XX PI Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
XX PI Williams PM, Wu TD, Zhang Z;
XX
XX DR WPI; 2003-354551/33.
XX DR N-PSDB; ACC49536.
XX
XX New antibodies against tumor-associated antigenic target polypeptide,
XX useful for treating or diagnosing tumors or cancers in mammals, e.g.
XX prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
XX carcinomas.
XX
XX Claim 2; Fig 100; 285pp; English.
XX
XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target
XX (TAT) proteins given in ABP97175 to ABP97234. The present invention
XX describes an isolated antibody that binds to a polypeptide having at
XX least 80 % sequence identity to any of the 60 150-800 residue amino acid
XX sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
XX its associated signal peptide, encoded by any of the 60 2000-3000 base
XX pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
XX cytostatic activity. The antibody can be used for treating or diagnosing
XX tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
XX cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
XX cell carcinomas, or thyroid cancer
XX
XX Sequence 252 AA;
XX
XX Query Match 100.0%; Score 252; DB 6; Length 252;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-233;
XX Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPFISHSQRRREDALSSGCLMPS 60
XX Db 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPFISHSQRRREDALSSGCLMPS 60
XX
XX QY 61 ESTVSGNGIPEPQVYAPPRPTDLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
XX Db 61 ESTVSGNGIPEPQVYAPPRPTDLAVPPPAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
XX
XX QY 121 BEPPPYQGCTQLQRPDPEQQLNRESVRAPNRTTFDSLDMSARLGCGPCPPSSNSGIS 180
XX Db 121 BEPPPYQGCTQLQRPDPEQQLNRESVRAPNRTTFDSLDMSARLGCGPCPPSSNSGIS 180
XX
XX QY 181 ATCYGSGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
XX Db 181 ATCYGSGGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
XX
XX QY 241 SKEKDQKQHPL 252
XX Db 241 SKEKDQKQHPL 252

```



```

Db      241  SKEKDQKQGHPL 252

RESULT 6
ADB75589
ID      ADB75589 standard; protein; 252 AA.
XX
AC      ADB75589;
XX
DT      04-DEC-2003 (first entry)
DE      Prostate cancer marker protein.
XX
KW      Prostate; cancer; cytostatic; gene therapy; marker.
XX
OS      Homo sapiens.
XX
PN      WO2003009814-A2.
XX
PD      06-FEB-2003.
XX
PF      25-JUL-2002; 2002WO-US023913.
XX
PR      25-JUL-2001; 2001US-0307982P.
PR      22-AUG-2001; 2001US-0314356P.
PR      25-SEP-2001; 2001US-0325020P.
PR      12-DEC-2001; 2001US-0341746P.
PR      05-MAR-2002; 2002US-0362158P.
XX
PA      (MILL-) MILLENNIUM PHARM INC.
XX
PI      Schlegel R, Monahan JB, Endege WO, Gannavarapu M, Gorbacheva B;
PI      Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX
DR      WPI; 2003-248033/24.
XX
XX      New nucleic acid molecule, useful for diagnosing or treating prostate
XX      cancer.
XX
PS      Disclosure; SEQ ID NO 413; 99pp; English.
XX
XX      The invention relates to newly discovered cancer markers associated with
XX      the cancerous state of prostate cells. Also disclosed is a method of
XX      assessing whether a patient is afflicted with prostate cancer. The method
XX      of the invention involves assessing whether a patient is afflicted with
XX      prostate cancer by comparing the level of expression of a marker in a
XX      patient sample and the normal level of expression of the marker in a
XX      control non-prostate cancer sample, where a significant increase in the
XX      level of expression of the marker in the patient sample and the normal
XX      level indicates that the patient is afflicted with prostate cancer.
XX      Nucleic acids of the invention are useful for diagnosing or treating
XX      prostate cancer, and may be useful in gene therapy. Sequences given in
XX      ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
XX      data for this patent did not form part of the printed specification, but
XX      was obtained in electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 252 AA;

Query Match      100.0%; Score 252; DB 7; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.8e-233;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAELFVQIIIIIVVMVMVMTCLLSHYKLSARSFISRHSGQRREDALSSEGCLWPS 60
DB      1  MAELFVQIIIIIVVMVMVMTCLLSHYKLSARSFISRHSGQRREDALSSEGCLWPS 60
QY      61  ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
DB      61  ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
QY      121  BEPPPYQGPCTQLQRLDPEQQLNRESVRAPNRTIFDSDLMDSARLGCPSPSSNSGIS 180

```

```

Db      121  BEPPPYQGPCTQLQRLDPEQQLNRESVRAPNRTIFDSDLMDSARLGCPSPSSNSGIS 180
QY      181  ATCYGSGRMGEGPPPTYSVICHYPGSSFOQOQSGSPSLLECTRLHHTHIAPLSAAIW 240
DB      181  ATCYGSGRMGEGPPPTYSVICHYPGSSFOQOQSGSPSLLECTRLHHTHIAPLSAAIW 240
QY      241  SKEKDQKQGHPL 252
DB      241  SKEKDQKQGHPL 252

RESULT 7
ADC37325
ID      ADC37325 standard; protein; 252 AA.
XX
AC      ADC37325;
XX
DT      18-DEC-2003 (first entry)
DE      Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 158.
XX
KW      Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW      cancer; infectious disease; bone disease; AIDS;
KW      neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW      Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW      Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
XX
OS      Homo sapiens.
XX
PN      WO2003048202-A2.
XX
PD      12-JUN-2003.
XX
PF      03-DEC-2002; 2002WO-JP012644.
XX
PR      03-DEC-2001; 2001JP-00368692.
PR      05-DEC-2001; 2001US-0035829P.
PR      03-OCT-2002; 2002JP-00291302.
PR      04-OCT-2002; 2002US-0415769P.
XX
PA      (ASAH ) ASahi KASEI KK.
XX
PI      Matsuda A, Muramatsu S;
XX
DR      WPI; 2003-505282/47.
DR      N-PSDB; ADC37324.
XX
XX      New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX      useful for treating inflammation, autoimmune diseases, cancers,
XX      infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX      ischemic disorders.
XX
PS      Claim 1; SEQ ID NO 158; 938pp; English.
XX
XX      The present invention relates to novel proteins and their coding
XX      sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
XX      kappaB). The proteins and their coding sequences are useful for treating
XX      a disease associated with NF-kappaB activation, such as inflammation,
XX      autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
XX      neurodegenerative diseases, or ischaemic disorders.
XX
SQ      Sequence 252 AA;

Query Match      100.0%; Score 252; DB 7; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.8e-233;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAELFVQIIIIIVVMVMVMTCLLSHYKLSARSFISRHSGQRREDALSSEGCLWPS 60
DB      1  MAELFVQIIIIIVVMVMVMTCLLSHYKLSARSFISRHSGQRREDALSSEGCLWPS 60
QY      61  ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
DB      61  ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
QY      121  BEPPPYQGPCTQLQRLDPEQQLNRESVRAPNRTIFDSDLMDSARLGCPSPSSNSGIS 180

```

Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
 QY 121 BEPPYQGPCTQLQDPEQQLINRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
 Db 121 BEPPYQGPCTQLQDPEQQLINRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
 QY 181 ATCYSGGGRMEGPPPTYSVIGHYFGSSFHQOQSSGPPSLLEGTRLLHHTHIAPLESAAIW 240
 Db 181 ATCYSGGGRMEGPPPTYSVIGHYFGSSFHQOQSSGPPSLLEGTRLLHHTHIAPLESAAIW 240
 QY 241 SKEKQKQKHPL 252
 Db 241 SKEKQKQKHPL 252

RESULT 8

AAE39589
 ID AAE39589 standard; protein; 252 AA.

XX AC AAE39589;

XX DT 18-DEC-2003 (first entry)

XX DE Human androgen-regulated gene (ARG), PMEPAL protein.

XX KW Human; androgen-regulated gene; ARG; PMEPAL; prostate cancer; chromosome 20q13.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Domain 9..25

XX FT /note= "Transmembrane domain"

XX FN US6566130-B1.

XX XX 20-MAY-2003.

XX XX 26-JAN-2001; 2001US-00769482.

XX PR 28-JAN-2000; 2000US-0178772P.

XX PR 31-JAN-2000; 2000US-0179045P.

XX PA (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.

XX PI Srivastava S, Moul JW, Xu LL, Segawa T;

XX DR WPI; 2003-719644/68.

XX DR N-PSDB; AAD60105.

XX PT Novel isolated androgen-regulated gene designated as PMEPAL useful for selecting primers and probes for detecting prostate cancer cells in biological samples by nucleic acid amplification techniques.

XX PS Claim 1; Col 29-30; 58pp; English.

XX CC The invention relates to an isolated androgen-regulated gene (ARG) designated as PMEPAL. The invention is useful for selecting primers and probes for detecting prostate cancer cells in a biological sample by using nucleic acid amplification techniques. The present sequence is human PMEPAL protein. PMEPAL gene chromosome 20q13

XX SQ Sequence 252 AA;

Query Match

Best Local Similarity 100.0%; Score 252; DB 7; Length 252;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELBFVQIIIVVWVWVITCLSHYKLSARSFISRHQGRREDALSSEGCCLWPS 60

Db 1 MAELBFVQIIIVVWVWVITCLSHYKLSARSFISRHQGRREDALSSEGCCLWPS 60

QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
 QY 121 BEPPYQGPCTQLQDPEQQLINRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
 Db 121 BEPPYQGPCTQLQDPEQQLINRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
 QY 181 ATCYSGGGRMEGPPPTYSVIGHYFGSSFHQOQSSGPPSLLEGTRLLHHTHIAPLESAAIW 240
 Db 181 ATCYSGGGRMEGPPPTYSVIGHYFGSSFHQOQSSGPPSLLEGTRLLHHTHIAPLESAAIW 240
 QY 241 SKEKQKQKHPL 252
 Db 241 SKEKQKQKHPL 252

RESULT 9

ABP75660

ID ABP75660 standard; protein; 285 AA.

XX AC ABP75660;

XX DT 10-FEB-2003 (first entry)

XX DE Human secretory polypeptide SPTM SEQ ID NO 844.

XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia; asthma; Crohn's disease; neurological disorder; epilepsy; cancer; Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease; multiple sclerosis; Parkinson's disease; cell proliferative disorder; anti-inflammatory; immunosuppressive; neuroprotective; nontropic; neuroleptic; anticonvulsant; cytotatic; antiparkinsonian; anxiolytic; antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus; secretory polynucleotide; secretory protein.

XX OS Homo sapiens.

XX PN WO200283876-A2.

XX PD 24-OCT-2002.

XX PF 27-MAR-2002; 2002WO-US009921.

XX PR 29-MAR-2001; 2001US-0280067P.

XX PR 29-MAR-2001; 2001US-0280068P.

XX PR 16-MAY-2001; 2001US-0291280P.

XX PR 17-MAY-2001; 2001US-0291829P.

XX PR 17-MAY-2001; 2001US-0291849P.

XX PR 19-JUN-2001; 2001US-0299428P.

XX PR 20-JUN-2001; 2001US-0299776P.

XX PR 20-JUN-2001; 2001US-0300001P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;

XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

XX PI Flores V, Matwaba R, Lo A, Lan RY, Urashka ME;

XX DR WPI; 2003-075543/07.

XX DR N-PSDB; ABZ36103.

XX PT New human secretory proteins and polynucleotides, useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell proliferations or cancers.

XX PS Claim 27; SEQ ID NO 844; 458pp + Sequence Listing; English.

XX CC The invention relates to a secretory polynucleotide (designated sptm) comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a naturally occurring polynucleotide sequence at least 90 % identical to

the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SPM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). The present sequence is one of the SPM proteins of the invention (ABP75384-ABP75962). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 285 AA;

Query Match 100.0%; Score 252; DB 6; Length 285;
Best Local Similarity 100.0%; Pred. No. 6.4e-233;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSPISRHSQGRREDALSSGCLWPS 60
Db 34 MAELFVQIIIVVMVMVVITCLSHYKLSARSPISRHSQGRREDALSSGCLWPS 93

Qy 61 ESTVSGNGIPEPQVYAPPRDRLAVPPPAQRERFHRFQPTYPVQLQHEIDLPTTISLSDG 120
Db 94 ESTVSGNGIPEPQVYAPPRDRLAVPPPAQRERFHRFQPTYPVQLQHEIDLPTTISLSDG 153

Qy 121 BEPPYQGPCTQLQRPDPEQQLNRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNGIS 180
Db 154 BEPPYQGPCTQLQRPDPEQQLNRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNGIS 213

Qy 181 ATCYGSGRMGEGPPPTTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTIAPLESAAIW 240
Db 214 ATCYGSGRMGEGPPPTTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTIAPLESAAIW 273

Qy 241 SKEDKQKGHPL 252
Db 274 SKEDKQKGHPL 285

RESULT 10
AAU78231
ID AAU78231 standard; protein; 287 AA.
XX AC AAU78231;
XX DT 05-JUN-2002 (first entry)
XX DB Human Mechanically Induced Vascular Receptor 1, MIVR-1.
XX KW Human; MIVR-1; Mechanically Induced Vascular; receptor; cytostatic;
KW cardiant; cerebroprotective; antiarteriosclerotic; cardiac cell;
KW anti-apoptotic; vascular endothelial cell; cardiac hypertrophy;
KW myocardial infarction; stroke; arteriosclerosis; heart failure.
XX OS Homo sapiens.
XX FN WO200216416-A2.
XX PD 28-FEB-2002.
XX PP 21-AUG-2001; 2001WO-US026089.
XX PR 22-AUG-2000; 2000US-0227159P.
XX XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX PA (PFIZ) PFIZER INC.

PI Lee RT, Landschulz KT, Kennedy SP, Thompson JP, Turi TG;
XX WPI; 2002-280912/32.
DR N-PSDB; ABK12137.
XX PT Novel nucleic acid molecule encoding Mechanically Induced Vascular
PT Receptor-1 polypeptide, useful for treating cardiovascular diseases.
XX Claim 14; Page 88-89; 105pp; English.
PS The invention relates to an isolated nucleic acid molecule encoding a
XX Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having the
CC cardiac cell anti-apoptotic activity and fragments of it provided they
CC are not identical to Genbank sequences AI761441.1, AI594390, NM_004338
CC and AQ177461. Also included are expression vectors, host cells, the MIVR-
CC 1 polypeptide, MIVR-1 binding peptides, modulators of MIVR-1, contacting
CC a molecule having cardiac cell anti-apoptotic activity with a candidate
CC agent, where the molecule is a nucleic acid molecule comprising MIVR-1,
CC IEX-1, VDUP-1, BTG-2 and TIS-11d or its expression product, determining
CC if the anti-apoptotic activity is modulated and thereby identifying a
CC modulator. The cardiac cell anti-apoptotic molecules and nucleic acids
CC of the invention are useful for treating, diagnosing and monitoring
CC progression of such diseases and disorders as characterised by increased
CC apoptotic cell-death of vascular endothelial cells e.g. cardiac
CC hypertrophy, myocardial infarction, stroke, arteriosclerosis and heart
XX failure. The present sequence represents human MIVR-1

Sequence 287 AA;

Query Match 99.2%; Score 250; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 5.3e-231;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ELEFVQIIIVVMVMVVITCLSHYKLSARSPISRHSQGRREDALSSGCLWPS 62
Db 38 ELEFVQIIIVVMVMVVITCLSHYKLSARSPISRHSQGRREDALSSGCLWPS 97

Qy 63 TVSGNGIPEPQVYAPPRDRLAVPPPAQRERFHRFQPTYPVQLQHEIDLPTTISLSDG 122
Db 98 TVSGNGIPEPQVYAPPRDRLAVPPPAQRERFHRFQPTYPVQLQHEIDLPTTISLSDG 157

Qy 123 PPTVQGPCTQLQRPDPEQQLNRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNGISAT 182
Db 158 PPTVQGPCTQLQRPDPEQQLNRESVRAPNRTIFDSDLMDSARLGCGPCPPSSNGISAT 217

Qy 183 CYGSGRMGEGPPPTTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTIAPLESAAIW 242
Db 218 CYGSGRMGEGPPPTTSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHTIAPLESAAIW 277

Qy 243 EKDKQKGHPL 252
Db 278 EKDKQKGHPL 287

RESULT 11
ABP7234
ID ABP7234 standard; protein; 287 AA.
XX AC ABP7234;
XX DT 01-JUL-2003 (first entry)
XX DE Tumour-associated antigenic target protein TAT379 cDNA SEQ ID NO:120.
XX KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
KW cancer.
XX OS Homo sapiens.
XX FN WO20003024392-A2.
XX PD 27-MAR-2003.

PP 11-SEP-2002; 2002WO-US028859.
 XX 18-SEP-2001; 2001US-0323268P.
 PR 19-OCT-2001; 2001US-0339227P.
 PR 07-NOV-2001; 2001US-0336827P.
 PR 20-NOV-2001; 2001US-0331906P.
 PR 02-JAN-2002; 2002US-0345444P.
 PR 03-APR-2002; 2002US-0369724P.
 PR 19-AUG-2002; 2002US-0404809P.
 XX (GETH) GENENTECH INC.
 XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
 PI Williams PM, Wu TD, Zhang Z;
 XX WPI; 2003-354551/33.
 DR N-PSDB; ACC49552.
 XX New antibodies against tumor-associated antigenic target polypeptide,
 PT useful for treating or diagnosing tumors or cancers in mammals, e.g.
 PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
 PT carcinomas.
 XX Claim 2; Fig 120; 285pp; English.
 PS ACC49493 to ACC49552 encode the human tumour-associated antigenic target
 CC (TAT) proteins given in ABP97175 to ABP97234. The present invention
 CC describes an isolated antibody that binds to a polypeptide having at
 CC least 80 % sequence identity to any of the 60 150-800 residue amino acid
 CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
 CC its associated signal peptide, encoded by any of the 60 2000-3000 base
 CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
 CC cytostatic activity. The antibody can be used for treating or diagnosing
 CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
 CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
 CC cell carcinomas, or thyroid cancer
 XX Sequence 287 AA;
 SQ
 Query Match 99.2%; Score 250; DB 6; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.3e-231;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 ELEFVQIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPSES 62
 Db 38 ELEFVQIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPSES 97
 QY 63 TVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLOHEIDLPTTISLSDGE 122
 Db 98 TVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLOHEIDLPTTISLSDGE 157
 QY 123 PPYPQGPCTQLQRLDPQQLNRESVRAPNRTIFDSLDMSARLGCPGPSNSGISAT 182
 Db 158 PPYPQGPCTQLQRLDPQQLNRESVRAPNRTIFDSLDMSARLGCPGPSNSGISAT 217
 QY 183 CYSGGGRMGPPPTYSVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 242
 Db 218 CYSGGGRMGPPPTYSVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 277
 QY 243 EKDKQKGHPL 252
 Db 278 EKDKQKGHPL 287
 RESULT 12
 ADC37327
 ID ADC37327 standard; protein; 287 AA.
 XX ADC37327;
 AC ADC37327;
 DT 18-DEC-2003 (first entry)
 XX Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 160.

XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
 OS Homo sapiens.
 XX WO2003048202-A2.
 XX 12-JUN-2003.
 XX 03-DEC-2002; 2002WO-JP012644.
 XX 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX (ASAH) ASAH KASBI KK.
 XX Matsuda A, Muramatsu S;
 PI WPI; 2003-505282/47.
 DR N-PSDB; ADC37326.
 XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX Claim 1; SEQ ID NO 160; 938pp; English.
 PS The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX Sequence 287 AA;
 SQ
 Query Match 99.2%; Score 250; DB 7; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.3e-231;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 ELEFVQIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPSES 62
 Db 38 ELEFVQIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPSES 97
 QY 63 TVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLOHEIDLPTTISLSDGE 122
 Db 98 TVSGNGIPEPQVYAPPRPTDRLAVPPPAQRERFHRFQPTYPYLOHEIDLPTTISLSDGE 157
 QY 123 PPYPQGPCTQLQRLDPQQLNRESVRAPNRTIFDSLDMSARLGCPGPSNSGISAT 182
 Db 158 PPYPQGPCTQLQRLDPQQLNRESVRAPNRTIFDSLDMSARLGCPGPSNSGISAT 217
 QY 183 CYSGGGRMGPPPTYSVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 242
 Db 218 CYSGGGRMGPPPTYSVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 277
 QY 243 EKDKQKGHPL 252
 Db 278 EKDKQKGHPL 287
 RESULT 13
 AAE39591
 ID AAE39591 standard; protein; 249 AA.
 XX AAE39591;
 AC AAE39591;

XX 18-DEC-2003 (first entry)
 XX Human androgen-regulated gene (ARG), PMEPAL protein fragment.
 DE Human; androgen-regulated gene; ARG; PMEPAL; prostate cancer; chromosome 20q13.
 KW Homo sapiens.
 OS US6566130-B1.
 XX 20-MAY-2003.
 XX 26-JAN-2001; 2001US-00769482.
 XX 28-JAN-2000; 2000US-0178772P.
 PR 31-JAN-2000; 2000US-0179045P.
 XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
 PA Srivastava S, Moul JW, Xu LL, Segawa T;
 PI WPI; 2003-719644/68.
 DR Novel isolated androgen-regulated gene designated as PMEPAL useful for
 PT selecting primers and probes for detecting prostate cancer cells in
 PT biological samples by nucleic acid amplification techniques.
 XX Example 7; Col 31-32; 59pp; English.
 XX The invention relates to an isolated androgen-regulated gene (ARG)
 CC designated as PMEPAL. The invention is useful for selecting primers and
 CC probes for detecting prostate cancer cells in a biological sample by
 CC using nucleic acid amplification techniques. The present sequence is
 CC human PMEPAL protein fragment. PMEPAL gene chromosome 20q13
 XX Sequence 249 AA;
 SQ
 Query Match 98.8%; Score 249; DB 7; Length 249;
 Best Local Similarity 100.0%; Pred. No. 4.3e-230;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALEFVQIIIVVMVWVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPSE 61
 DB 1 ALEFVQIIIVVMVWVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPSE 60
 QY 62 STVSGNGIPEQVYAPPRTDLAVPPFAQRERFHRFQPTTYLQHEIDLPTTISLSDGE 121
 DB 61 STVSGNGIPEQVYAPPRTDLAVPPFAQRERFHRFQPTTYLQHEIDLPTTISLSDGE 120
 QY 122 EPPYQGPCTQLRDPQOOLELNRESVRAPPNRTIFDSLDMSARLGGCPSPSSNGISA 181
 DB 121 EPPYQGPCTQLRDPQOOLELNRESVRAPPNRTIFDSLDMSARLGGCPSPSSNGISA 180
 QY 182 TCYGGGMEGPPPTYSVIGHYRGSSFOHQSSGPPSILECTRLHHTHIAPLESAAIWS 241
 DB 181 TCYGGGMEGPPPTYSVIGHYRGSSFOHQSSGPPSILECTRLHHTHIAPLESAAIWS 240
 QY 242 KEKDKQKH 250
 DB 241 KEKDKQKH 249
 RESULT 14
 AAB18461
 ID AAB18461 standard; protein; 252 AA.
 XX AAB18461;
 AC AAB18461;
 XX 15-JAN-2001 (first entry)
 DT A human TANGO 261 polypeptide clone.
 DE

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
 KW cellular proliferation; cellular differentiation; cellular adhesion;
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
 KW intestinal disorder; spleen associated disease; renal disorder;
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
 KW brain herniation; iatrogenic disease; inflammation; meningitis;
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.
 XX Homo sapiens.
 XX WO200052022-A1.
 XX 08-SEP-2000.
 XX 01-MAR-2000; 2000WO-US005226.
 XX 01-MAR-1999; 99US-0122458P.
 PR (MILL-) MILLENNIUM PHARM INC.
 XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
 PI WPI; 2000-579269/54.
 DR N-PSDB; AAA75163.
 XX Novel human and murine secreted proteins designated TANGO 216, 261, 262,
 PT 266 and 267 useful as modulating agents of cellular processes, e.g. for
 PT treating cancer.
 PT Disclosure; Page; 175pp; English.
 XX AAB18461-63 represent human TANGO 261 proteins. The specification also
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
 CC polypeptides can be used to modulate cellular proliferation, modulate
 CC cellular differentiation and/or modulate cellular adhesion. The proteins
 CC can be used to treat any von Willebrand factor-associated disorder,
 CC regulate extracellular matrix structuring, cellular adhesion, and cell
 CC trafficking and/or migration, modulate cellular interactions, modulate the
 CC cell adhesion in proliferative disorders, such as cancer, modulate the
 CC proliferation, differentiation, and/or function of cells that appear in
 CC the bone marrow, and leukocytes, treat bone marrow, blood and
 CC hematopoietic associated diseases and disorders, atelectasis, pulmonary
 CC congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and
 CC bronchiectasis, intestinal disorders, spleen associated diseases,
 CC modulate renal disorders, treat cardiovascular disorders such as ischemic
 CC heart disease, modulate the proliferation, differentiation, and/or
 CC function of bone and cartilage cells and to treat bone and/or cartilage
 CC associated diseases or disorder. They may also be used to treat disorders
 CC associated with the ovaries, and cerebral oedema, hydrocephalus, brain
 CC herniations, iatrogenic disease, inflammations, bacterial and viral
 CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
 CC disease, multiple sclerosis, brain cancers, hydrocephalus and
 CC encephalitis, and treat hepatic disorders. note: the present sequence
 CC does not appear in the specification; it was created using information
 CC provided
 XX Sequence 252 AA;
 SQ
 Query Match 98.8%; Score 249; DB 3; Length 252;
 Best Local Similarity 100.0%; Pred. No. 4.3e-230;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 LEFVQIIIVVMVWVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPSEST 63
 DB 4 LEFVQIIIVVMVWVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPSEST 63
 QY 64 VSGNGIPEQVYAPPRTDLAVPPFAQRERFHRFQPTTYLQHEIDLPTTISLSDGE 123
 DB 64 VSGNGIPEQVYAPPRTDLAVPPFAQRERFHRFQPTTYLQHEIDLPTTISLSDGE 123

QY 124 PPVQGGCTTQLRDEPQQLNRESVRAPPNRTTFDSLDMSARLGPGPCPPSSNSGISATC 183
 |||||
 Db 124 PPVQGGCTTQLRDEPQQLNRESVRAPPNRTTFDSLDMSARLGPGPCPPSSNSGISATC 183
 |||||
 QY 184 YGSGGRMEGPPPTYSYEVIGHYPGSSFOHQSSGPPSLLEGTTLHHTHIAPIESAAIWSKE 243
 |||||
 Db 184 YGSGGRMEGPPPTYSYEVIGHYPGSSFOHQSSGPPSLLEGTTLHHTHIAPIESAAIWSKE 243
 |||||
 QY 244 KDKQKGHPL 252
 |||||
 Db 244 KDKQKGHPL 252
 |||||

RESULT 15

AAM38712
 ID AAM38712 standard; protein; 237 AA.

XX

AC AAM38712;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 1857.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

OS

XX WO200153312-A1.

PN

XX 26-JUL-2001.

PD

XX 26-DEC-2000; 2000MO-US034263.

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00489725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR N-PSDB; AAI57868.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Example 3; SEQ ID NO 1857; 10078pp; English.

PS The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 237 AA;

Query Match 94.0%; Score 237; DB 4; Length 237;

Best Local Similarity 100.0%; Pred. No. 1.3e-218;

Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMVMVVVITCLLSHYKLSARSFISRHSGRRRDLSSSECLWPFSESTVSGNGIPEPOVY 75

Db 1 MMVMVVVITCLLSHYKLSARSFISRHSGRRRDLSSSECLWPFSESTVSGNGIPEPOVY 60

QY 76 APPRPTDLRAVPPFAQRERFHRFPQTPYLYLQHEIDLPTTISLSDGEEPPPYQGCTQLQR 135

Db 61 APPRPTDLRAVPPFAQRERFHRFPQTPYLYLQHEIDLPTTISLSDGEEPPPYQGCTQLQR 120

QY 136 DPEQQLNRESVRAPPNRTTFDSLDMSARLGPGPCPPSSNSGISATCYGSGGRMEGPPPP 195

Db 121 DPEQQLNRESVRAPPNRTTFDSLDMSARLGPGPCPPSSNSGISATCYGSGGRMEGPPPP 180

QY 196 TYSEVIGHYPGSSFOHQSSGPPSLLEGTTLHHTHIAPIESAAIWSKEKDKQKGHPL 252

Db 181 TYSEVIGHYPGSSFOHQSSGPPSLLEGTTLHHTHIAPIESAAIWSKEKDKQKGHPL 237

Search completed: May 20, 2004, 15:34:47

Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 20, 2004, 15:33:51 ; Search time 23 Seconds
(without alignments)
565.641 Million cell updates/sec

Title: US-09-857-826b-17
Perfect score: 252
Sequence: 1 MAELFVQIIIVVMVMMVV.....PLESAAIWSKEKQKQHPL 252

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/2/iaa/5A.COMB.pap.*
2: /cgm2_6/ptodata/2/iaa/5B.COMB.pap.*
3: /cgm2_6/ptodata/2/iaa/6A.COMB.pap.*
4: /cgm2_6/ptodata/2/iaa/6B.COMB.pap.*
5: /cgm2_6/ptodata/2/iaa/PTCUS.COMB.pap.*
6: /cgm2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	252	100.0	252	4 US-09-769-482-3
2	249	98.8	249	4 US-09-769-482-11
3	35	13.9	244	4 US-09-769-482-12
4	35	13.9	288	4 US-09-091-952A-4
5	35	13.9	306	4 US-09-091-952A-3
6	14	5.6	69	4 US-09-621-976-6022
7	7	2.8	18	4 US-09-091-952A-5
8	7	2.8	269	4 US-09-134-000C-5214
9	7	2.8	272	4 US-09-252-991A-23410
10	7	2.8	285	4 US-09-149-476-470
11	7	2.8	289	3 US-08-872-979-6
12	7	2.8	326	4 US-09-107-532A-3665
13	7	2.8	369	4 US-09-252-991A-27769
14	7	2.8	407	4 US-09-252-991A-20436
15	7	2.8	448	4 US-09-518-657-2
16	7	2.8	450	4 US-09-518-657-4
17	7	2.8	540	4 US-09-513-057C-33
18	7	2.8	551	4 US-09-252-991A-26416
19	7	2.8	589	1 US-08-399-646-2
20	7	2.8	589	1 US-08-607-321-2
21	7	2.8	589	2 US-08-961-240-2
22	7	2.8	589	2 US-08-605-501-2
23	7	2.8	589	1 US-08-399-646-12
24	7	2.8	596	1 US-08-607-321-12
25	7	2.8	596	1 US-08-961-240-12
26	7	2.8	596	2 US-08-605-501-12
27	7	2.8	664	2 US-08-852-153-8

28	7	2.8	665	4	US-09-228-986-68	Sequence 68, Appl
29	6	2.4	21	4	US-09-288-143-178	Sequence 178, Appl
30	6	2.4	37	3	US-08-470-369-19	Sequence 19, Appl
31	6	2.4	37	6	5304637-19	Patent No. 5304637
32	6	2.4	46	3	US-08-905-223-347	Sequence 347, Appl
33	6	2.4	74	4	US-09-621-976-7191	Sequence 7191, Appl
34	6	2.4	86	1	US-08-390-858B-39	Sequence 39, Appl
35	6	2.4	95	4	US-09-309-572-8	Sequence 8, Appl
36	6	2.4	95	4	US-09-718-096-8	Sequence 8, Appl
37	6	2.4	99	4	US-09-621-976-4793	Sequence 4793, Appl
38	6	2.4	101	4	US-09-621-976-7124	Sequence 7124, Appl
39	6	2.4	102	4	US-09-289-198-308	Sequence 308, Appl
40	6	2.4	102	4	US-09-429-755-308	Sequence 308, Appl
41	6	2.4	103	1	US-08-376-296-10	Sequence 10, Appl
42	6	2.4	110	1	US-08-844-064-8	Sequence 8, Appl
43	6	2.4	110	3	US-09-009-433-8	Sequence 8, Appl
44	6	2.4	112	4	US-09-621-976-6250	Sequence 6250, Appl
45	6	2.4	120	4	US-09-615-192A-393	Sequence 393, Appl

ALIGNMENTS

RESULT 1
US-09-769-482-3
; Sequence 3, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: POYNUCLEOTIDE ARRAY
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-3

Query Match 100.0%; Score 252; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.4e-239;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAELFVQIIIVVMVMMVVVITCLSHYKLSARSPTSRSQGRRRRDLSSGCLWPS	60
Db	1	MAELFVQIIIVVMVMMVVVITCLSHYKLSARSPTSRSQGRRRRDLSSGCLWPS	60
Qy	61	ESTVSGNGIPQVYAPRPTDLAVPPAQRERPHRFQPTYPYLOHEIDLPTTISLSDG	120
Db	61	ESTVSGNGIPQVYAPRPTDLAVPPAQRERPHRFQPTYPYLOHEIDLPTTISLSDG	120
Qy	121	BEPPYQPCCTQLRDPEQLANRESVRAPNRTIFDSLDMSARLGGPCPPSSNGIS	180
Db	121	BEPPYQPCCTQLRDPEQLANRESVRAPNRTIFDSLDMSARLGGPCPPSSNGIS	180
Qy	181	ATCYGSGRMGPPTTYSVICHYFGSSFOHQSSGPPSLLEGTLLHHTHIAPLSAAIW	240
Db	181	ATCYGSGRMGPPTTYSVICHYFGSSFOHQSSGPPSLLEGTLLHHTHIAPLSAAIW	240
Qy	241	SKKDKQKQHPL	252
Db	241	SKKDKQKQHPL	252

filed 1-26-07

evaluated 1-28-07

```
RESULT 2
US-09-769-482-11
; Sequence 11, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: POYNUCLEOTIDE ARRAY
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-11

Query Match      98.8%; Score 249; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.1e-236;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  ALEFVQIIIVVMVMVVTCLSHYKLSARSFISHSQGRREDALSSGCLWPSE 61
DB      1  ALEFVQIIIVVMVMVVTCLSHYKLSARSFISHSQGRREDALSSGCLWPSE 60

QY      62  STVSGNGIPEQVAPPRTDRLAVPPFAQRERFHPQTPYVLOHEIDLPTTISLDCGE 121
DB      61  STVSGNGIPEQVAPPRTDRLAVPPFAQRERFHPQTPYVLOHEIDLPTTISLDCGE 120

QY      122  EPPYQGPCTQLRDPEQOELNRESVRAPPNTIFDSDLMDSARLGCGCPSSNSGISA 181
DB      121  EPPYQGPCTQLRDPEQOELNRESVRAPPNTIFDSDLMDSARLGCGCPSSNSGISA 180

QY      182  TCVSGGRMEGPPPTYSVIGHYVPGSSFOHQSSGPPSLLEGTRLHETHIAPLESAAIWS 241
DB      181  TCVSGGRMEGPPPTYSVIGHYVPGSSFOHQSSGPPSLLEGTRLHETHIAPLESAAIWS 240

QY      242  KKKKQKQH 250
DB      241  KKKKQKQH 249

RESULT 3
US-09-769-482-12
; Sequence 12, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: POYNUCLEOTIDE ARRAY
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
```

```
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-12

Query Match      13.9%; Score 35; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      106  QHEIDLPTTISLSDGHEPPPYQGPCTQLRDPEQQ 140
DB      103  QHEIDLPTTISLSDGHEPPPYQGPCTQLRDPEQQ 137

RESULT 4
US-09-091-952A-4
; Sequence 4, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; APPLICANT: Gershon, Elliot S.
; APPLICANT: Badner, Judith A.
; APPLICANT: Goldin, Lynn R.
; APPLICANT: Berrettini, Wade H.
; APPLICANT: Yoshikawa, Takeo
; APPLICANT: Sanders, Alan R.
; APPLICANT: Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..288
; OTHER INFORMATION: Clone 22 isoform 2
; alternatively spliced
; protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
```


US-09-091-952A-4

Query Match 13.9%; Score 35; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 2.4e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 QHEIDLPTTSLSDGEBPPPYQGCTQLQLRDPEQQ 140
Db 145 QHEIDLPTTSLSDGEBPPPYQGCTQLQLRDPEQQ 179

RESULT 5

US-09-091-952A-3
; Sequence 3, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.

; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.

; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/091.952A

; FILING DATE: 19-Apr-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/029,278

; FILING DATE: 28-OCT-1996

; APPLICATION NUMBER: PCT/US97/19381

; FILING DATE: 28-OCT-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, Timothy L.

; REGISTRATION NUMBER: 35,367

; REFERENCE/DOCKET NUMBER: 015280-297100US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 306 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 1...306

; OTHER INFORMATION: Clone 22 isoform 1,

; unspliced protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-091-952A-3

Query Match 13.9%; Score 35; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.5e-26;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 QHEIDLPTTSLSDGEBPPPYQGCTQLQLRDPEQQ 140
Db 163 QHEIDLPTTSLSDGEBPPPYQGCTQLQLRDPEQQ 197

RESULT 6

US-09-621-976-6022
; Sequence 6022, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621.976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 6022

; LENGTH: 69

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-621-976-6022

Query Match

Best Local Similarity 5.6%; Score 14; DB 4; Length 69;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 GGCPSPSSNSGISA 181

Db 4 GGCPSPSSNSGISA 17

RESULT 7

US-09-091-952A-5

; Sequence 5, Application US/09091952A

; Patent No. 6458532

; GENERAL INFORMATION:

; APPLICANT: Detera-Wadleigh, Sevilla D.

; Gershon, Elliot S.

; Badner, Judith A.

; Goldin, Lynn R.

; Berrettini, Wade H.

; Yoshikawa, Takeo

; Sanders, Alan R.

; Esterling, Lisa E.

; TITLE OF INVENTION: Chromosomal Markers and Diagnostic

; Tests for Manic-Depressive Illness

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/091.952A

; FILING DATE: 19-Apr-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/029,278

; FILING DATE: 28-OCT-1996

; APPLICATION NUMBER: PCT/US97/19381

; FILING DATE: 28-OCT-1997

ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-2971000S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: -
LOCATION: 1...18
OTHER INFORMATION: alternatively spliced portion
lacking from isoform 2
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-091-952A-5

Query Match 2.8%; Score 7; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 EGCLWPS 60
|||||||
DB 1 EGCLWPS 7

RESULT 8

US-09-134-000C-5214
; Sequence 5214, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5214
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5214

Query Match 2.8%; Score 7; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GPCPPPS 174
|||||||
DB 84 GPCPPPS 90

RESULT 9

US-09-252-991A-23410
; Sequence 23410, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23410
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23410

Query Match 2.8%; Score 7; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 RLGGPCP 172
|||||||
DB 33 RLGGPCP 39

RESULT 10

US-09-149-476-470
; Sequence 470, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23

1	EARLIER	FILING DATE:	1997-08-22
2	EARLIER	APPLICATION NUMBER:	60/056,911
3	EARLIER	FILING DATE:	1997-08-22
4	EARLIER	APPLICATION NUMBER:	60/056,636
5	EARLIER	FILING DATE:	1997-08-22
6	EARLIER	APPLICATION NUMBER:	60/056,874
7	EARLIER	FILING DATE:	1997-08-22
8	EARLIER	APPLICATION NUMBER:	60/056,910
9	EARLIER	FILING DATE:	1997-08-22
10	EARLIER	APPLICATION NUMBER:	60/056,864
11	EARLIER	FILING DATE:	1997-08-22
12	EARLIER	APPLICATION NUMBER:	60/056,631
13	EARLIER	FILING DATE:	1997-08-22
14	EARLIER	APPLICATION NUMBER:	60/056,845
15	EARLIER	FILING DATE:	1997-08-22
16	EARLIER	APPLICATION NUMBER:	60/056,892
17	EARLIER	FILING DATE:	1997-08-22
18	EARLIER	APPLICATION NUMBER:	60/057,761
19	EARLIER	FILING DATE:	1997-08-22
20	EARLIER	APPLICATION NUMBER:	60/047,595
21	EARLIER	FILING DATE:	1997-05-23
22	EARLIER	APPLICATION NUMBER:	60/047,599
23	EARLIER	FILING DATE:	1997-05-23
24	EARLIER	APPLICATION NUMBER:	60/047,588
25	EARLIER	FILING DATE:	1997-05-23
26	EARLIER	APPLICATION NUMBER:	60/047,585
27	EARLIER	FILING DATE:	1997-05-23
28	EARLIER	APPLICATION NUMBER:	60/047,586
29	EARLIER	FILING DATE:	1997-05-23
30	EARLIER	APPLICATION NUMBER:	60/047,590
31	EARLIER	FILING DATE:	1997-05-23
32	EARLIER	APPLICATION NUMBER:	60/047,594
33	EARLIER	FILING DATE:	1997-05-23
34	EARLIER	APPLICATION NUMBER:	60/047,589
35	EARLIER	FILING DATE:	1997-05-23
36	EARLIER	APPLICATION NUMBER:	60/047,593
37	EARLIER	FILING DATE:	1997-05-23
38	EARLIER	APPLICATION NUMBER:	60/047,614
39	EARLIER	FILING DATE:	1997-05-23
40	EARLIER	APPLICATION NUMBER:	60/043,578
41	EARLIER	FILING DATE:	1997-04-11
42	EARLIER	APPLICATION NUMBER:	60/043,576
43	EARLIER	FILING DATE:	1997-04-11
44	EARLIER	APPLICATION NUMBER:	60/047,501
45	EARLIER	FILING DATE:	1997-05-23
46	EARLIER	APPLICATION NUMBER:	60/043,670
47	EARLIER	FILING DATE:	1997-04-11
48	EARLIER	APPLICATION NUMBER:	60/056,632
49	EARLIER	FILING DATE:	1997-08-22
50	EARLIER	APPLICATION NUMBER:	60/056,664
51	EARLIER	FILING DATE:	1997-08-22
52	EARLIER	APPLICATION NUMBER:	60/056,876
53	EARLIER	FILING DATE:	1997-08-22
54	EARLIER	APPLICATION NUMBER:	60/056,881
55	EARLIER	FILING DATE:	1997-08-22
56	EARLIER	APPLICATION NUMBER:	60/056,909
57	EARLIER	FILING DATE:	1997-08-22
58	EARLIER	APPLICATION NUMBER:	60/056,875
59	EARLIER	FILING DATE:	1997-08-22
60	EARLIER	APPLICATION NUMBER:	60/056,862
61	EARLIER	FILING DATE:	1997-08-22
62	EARLIER	APPLICATION NUMBER:	60/056,887
63	EARLIER	FILING DATE:	1997-08-22
64	EARLIER	APPLICATION NUMBER:	60/056,908
65	EARLIER	FILING DATE:	1997-08-22
66	EARLIER	APPLICATION NUMBER:	60/048,964
67	EARLIER	FILING DATE:	1997-06-06
68	EARLIER	APPLICATION NUMBER:	60/057,650
69	EARLIER	FILING DATE:	1997-09-05
70	EARLIER	APPLICATION NUMBER:	60/056,884
71	EARLIER	FILING DATE:	1997-08-22
72	EARLIER	APPLICATION NUMBER:	60/057,669
73	EARLIER	FILING DATE:	1997-09-01

; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 2.8%; Score 7; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 SSGPPSL 220
DB 38 SSGPPSL 44

RESULT 11

US-08-872-979-6
; Sequence 6, Application US/08872979
; Patent No. 6074844
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,979
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0320 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 924268

US-08-872-979-6

Query Match 2.8%; Score 7; DB 3; Length 289;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IIIIVV 15
DB 265 IIIIVV 271

RESULT 12

US-107-532A-3665

; Sequence 3665, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3665:

SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...326

SEQUENCE DESCRIPTION: SEQ ID NO: 3665:

US-09-107-532A-3665

Query Match 2.8%; Score 7; DB 4; Length 326;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 MDSARLG 168
DB 93 MDSARLG 99

RESULT 13

US-09-252-991A-27769
; Sequence 27769, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27769
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27769

Query Match 2.8%; Score 7; DB 4; Length 369;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 HSQRRR 47
|||
Db 258 HSQRRR 264

RESULT 14

US-09-252-991A-20436
; Sequence 20436, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,786
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20436
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20436

Query Match 2.8%; Score 7; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 PRPTDRL 84
|||
Db 246 PRPTDRL 252

RESULT 15

US-09-518-657-2
; Sequence 2, Application US/09518657
; Patent No. 6335188
; GENERAL INFORMATION:
; APPLICANT: Schardl, Christopher L.
; APPLICANT: Wang, Jinghong
; TITLE OF INVENTION: Endophyte Ergot Alkaloid Synthetic Compounds, Compounds
; TITLE OF INVENTION: Which Encode Therefor and Related Methods
; FILE REFERENCE: P-1060
; CURRENT APPLICATION NUMBER: US/09/518,657
; CURRENT FILING DATE: 2000-03-03
; EARLIER APPLICATION NUMBER: 60/125,490
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Neotyphodium coenophialum
US-09-518-657-2

Query Match 2.8%; Score 7; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 IPEPQVY 75
|||
Db 346 IPEPQVY 352

Search completed: May 20, 2004, 15:37:22
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:36:21 ; Search time 48 Seconds
(without alignments)
1464.339 Million cell updates/sec

Title: US-09-857-826b-17
Perfect score: 252
Sequence: 1 MAELFVQIIIIIVVMVWV.....PLESAALWSKDKQKHPL 252

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1149313 seqs, 278921704 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252	100.0	252	10	US-09-796-753-56
2	252	100.0	252	10	US-09-821-812-3
3	252	100.0	252	14	US-10-205-823-413
4	252	100.0	252	14	US-10-241-220-100
5	252	100.0	252	14	US-10-241-220-101
6	252	100.0	252	14	US-10-301-822-209
7	252	100.0	252	14	US-10-390-045-3
8	252	100.0	252	16	US-10-434-479-3
9	250	99.2	287	9	US-09-934-249-2
10	250	99.2	287	14	US-10-241-220-120
11	250	99.2	287	15	US-10-235-027-128
12	249	98.8	249	14	US-10-390-045-11
13	249	98.8	249	16	US-10-434-479-11
14	181	71.8	241	14	US-10-000-256A-158
15	51	20.2	217	10	US-09-796-753-58

Sequence 68, Appl
Sequence 13, Appl
Sequence 12, Appl
Sequence 1836, Ap
Sequence 0, Appl
Sequence 1668, Ap
Sequence 0, Appl
Sequence 49175, A
Sequence 71739, A
Sequence 5, Appl
Sequence 35019, A
Sequence 37387, A
Sequence 34567, A
Sequence 192751, A
Sequence 210969, A
Sequence 45782, A
Sequence 216094, A
Sequence 8357, Ap
Sequence 34191, A
Sequence 33541, A
Sequence 14686, A
Sequence 48837, A
Sequence 206919, A
Sequence 651, App
Sequence 37593, A
Sequence 46070, A
Sequence 44, Appl
Sequence 53348, A

243 16 US-10-434-479-68
274 9 US-09-934-249-13
244 14 US-10-390-045-12
244 16 US-10-434-479-12
269 15 US-10-094-749-1836
288 14 US-10-251-598-4
306 12 US-10-087-192-1668
306 14 US-10-251-598-3
453 12 US-10-425-114-49175
585 12 US-10-425-114-71739
18 14 US-10-251-598-5
32 9 US-09-864-761-35019
37 9 US-09-864-761-37387
41 9 US-09-864-761-34567
64 12 US-10-424-599-192751
70 12 US-10-424-599-210969
72 9 US-09-864-761-45782
88 12 US-10-335-977-8356
89 12 US-10-424-599-216094
94 12 US-10-335-977-8357
105 9 US-09-864-761-34191
109 9 US-09-864-761-33541
119 14 US-10-156-761-14686
129 12 US-10-425-114-48837
136 12 US-10-424-599-206919
146 16 US-10-427-692-651
158 12 US-10-425-114-37993
205 12 US-10-425-114-46070
228 15 US-10-200-012-44
247 12 US-10-425-114-53348

ALIGNMENTS

RESULT 1
US-09-796-753-56
; Sequence 56, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23

```

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 56
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-56

```

```

Query Match      100.0%; Score 252; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIVVMMVVVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
DB 1 MALEFVQIIIVVMMVVVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60

QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

QY 121 BEPPYQGPCTQLQDRDPEQQLNRESVRAPNRTIFDSLDMSARLGCPPPSSNSGIS 180
DB 121 BEPPYQGPCTQLQDRDPEQQLNRESVRAPNRTIFDSLDMSARLGCPPPSSNSGIS 180

QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSGFHQOQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
DB 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSGFHQOQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

QY 241 SKEKDKQKHPL 252
DB 241 SKEKDKQKHPL 252

```

```

RESULT 2
US-09-821-812-3
; Sequence 3, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biayang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; FILE REFERENCE: P-15 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

US-09-821-812-3
Query Match      100.0%; Score 252; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIVVMMVVVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60
DB 1 MALEFVQIIIVVMMVVVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60

QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

QY 121 BEPPYQGPCTQLQDRDPEQQLNRESVRAPNRTIFDSLDMSARLGCPPPSSNSGIS 180
DB 121 BEPPYQGPCTQLQDRDPEQQLNRESVRAPNRTIFDSLDMSARLGCPPPSSNSGIS 180

QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSGFHQOQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
DB 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSGFHQOQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

QY 241 SKEKDKQKHPL 252
DB 241 SKEKDKQKHPL 252

```

```

RESULT 3
US-10-205-823-413
; Sequence 413, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John B.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wansley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 413
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-413

```

```

Query Match      100.0%; Score 252; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIVVMMVVVVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60

```

```

Db      1  MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPS 60
QY      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
Db      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
QY      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
Db      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
QY      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
Db      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
QY      241  SKEKDQKQGHPL 252
        |||||||
Db      241  SKEKDQKQGHPL 252
        |||||||

```

RESULT 4

```

US-10-241-220-100
; Sequence 100, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 100
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-100

```

```

Query Match      100.0%; Score 252; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPS 60
        |||||||
Db      1  MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPS 60
        |||||||
QY      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
Db      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
QY      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
Db      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
QY      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
Db      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
QY      241  SKEKDQKQGHPL 252
        |||||||
Db      241  SKEKDQKQGHPL 252
        |||||||

```

RESULT 5

```

US-10-241-220-101
; Sequence 101, Application US/10241220
; Publication No. US20030148408A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 101
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-101

```

```

Query Match      100.0%; Score 252; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.7e-231;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPS 60
        |||||||
Db      1  MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSGRRRDLSSGCLWPS 60
        |||||||
QY      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
Db      61  ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
        |||||||
QY      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
Db      121  BEPPPYQGCTQLQRDPEQQLNRSSVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
        |||||||
QY      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
Db      181  ATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
        |||||||
QY      241  SKEKDQKQGHPL 252
        |||||||
Db      241  SKEKDQKQGHPL 252
        |||||||

```

RESULT 6

```

US-10-301-822-209
; Sequence 209, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: WPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0

```


; SEQ ID NO 209
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-301-822-209

 Query Match 100.0%; Score 252; DB 14; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.7e-231;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60
 DB 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60

 QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
 DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

 QY 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180
 DB 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180

 QY 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
 DB 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

 QY 241 SKEKDQKQGHPL 252
 DB 241 SKEKDQKQGHPL 252

RESULT 7
 US-10-390-045-3
 ; Sequence 3, Application US/10390045
 ; Publication No. US20030170713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SRIVASTAVA, SHIV
 ; APPLICANT: MOUL, JUDD W.
 ; APPLICANT: SEGAWA, TAKEHIKO
 ; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
 ; FILE REFERENCE: 04995.0057-00000
 ; CURRENT APPLICATION NUMBER: US/10/390,045
 ; CURRENT FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: US/09/769,482
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 60/178,772
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: 60/179,045
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-390-045-3

Query Match 100.0%; Score 252; DB 14; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.7e-231;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60
 DB 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60

 QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
 DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

 QY 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180
 DB 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180

Db 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180
 QY 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
 Db 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
 QY 241 SKEKDQKQGHPL 252
 Db 241 SKEKDQKQGHPL 252

 RESULT 8
 US-10-434-479-3
 ; Sequence 3, Application US/10434479
 ; Publication No. US20040092469A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SRIVASTAVA, SHIV
 ; APPLICANT: MOUL, JUDD W.
 ; APPLICANT: XU, LINDA L.
 ; TITLE OF INVENTION: ANDROGEN-REGULATED PMP2A GENE AND POLYPEPTIDES
 ; FILE REFERENCE: 04995.0057-02000
 ; CURRENT APPLICATION NUMBER: US/10/434,479
 ; CURRENT FILING DATE: 2003-05-09
 ; PRIOR APPLICATION NUMBER: 10/390,045
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: 09/769,482
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 60/178,772
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: 60/179,045
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-434-479-3

Query Match 100.0%; Score 252; DB 16; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.7e-231;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60
 DB 1 MAELFVQIIIIIVVMVWVITCLSHYKLSARSPFISRHSGRRREDALSSGCLMPS 60

 QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120
 DB 61 ESTVSGNGIPEQVYAPRPTDRLAVPPPAQRERFHRFOPTYPYLQHEIDLPTTISLSDG 120

 QY 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180
 DB 121 BEPPYQGPCTQLQDRPEQQLNELNRESVRAPPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180

 QY 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
 DB 181 ATCYSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

 QY 241 SKEKDQKQGHPL 252
 DB 241 SKEKDQKQGHPL 252

 RESULT 9
 US-09-934-249-2
 ; Sequence 2, Application US/09934249
 ; Patent No. US20020115081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Richard T.
 ; APPLICANT: Landschulz, Katherine T.
 ; APPLICANT: Turi, Thomas G.
 ; APPLICANT: Thompson, John F.

```

; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/BRP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-934-249-2

```

Query Match 99.2%; Score 250; DB 9; Length 287;
 Best Local Similarity 100.0%; Pred. No. 1.5e-229;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 62
DB 38 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 97

QY 63 TVSGNGIPEPQVAPRPTDRLAVPPAQRERFHRPQPTPYLOHEIDLPTTISLSDGEE 122
DB 98 TVSGNGIPEPQVAPRPTDRLAVPPAQRERFHRPQPTPYLOHEIDLPTTISLSDGEE 157

QY 123 PPYPQGPCTQLRDPQQLLELNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 182
DB 158 PPYPQGPCTQLRDPQQLLELNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 217

QY 183 CYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 242
DB 218 CYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 277

QY 243 EKDKQKHPL 252
DB 278 EKDKQKHPL 287

```

```

RESULT 10
US-10-241-220-120
; Sequence 120, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakie, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 120
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-120

```

Query Match 99.2%; Score 250; DB 14; Length 287;
 Best Local Similarity 100.0%; Pred. No. 1.5e-229;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 62
DB 38 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 97

```

```

QY 63 TVSGNGIPEPQVAPRPTDRLAVPPAQRERFHRPQPTPYLOHEIDLPTTISLSDGEE 122
DB 98 TVSGNGIPEPQVAPRPTDRLAVPPAQRERFHRPQPTPYLOHEIDLPTTISLSDGEE 157

QY 123 PPYPQGPCTQLRDPQQLLELNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 182
DB 158 PPYPQGPCTQLRDPQQLLELNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGISAT 217

QY 183 CYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 242
DB 218 CYGSGRMEGPPPTYSVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLSAAIWSK 277

QY 243 EKDKQKHPL 252
DB 278 EKDKQKHPL 287

```

```

RESULT 11
US-10-295-027-128
; Sequence 128, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 128
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-128

```

Query Match 99.2%; Score 250; DB 15; Length 287;
 Best Local Similarity 100.0%; Pred. No. 1.5e-229;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 62
DB 38 ELEFVQIIIVVMVMMVITCLSHYKLSARSTFHSQRRREDALSSBGLWPSES 97

```

Db 38 ELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSE 97
Qy 63 TVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 122
Db 98 TVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 157
Qy 123 PYPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 182
Db 158 PYPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 217
Qy 183 CYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 242
Db 218 CYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 277
Qy 243 EKDKQKQGHPL 252
Db 278 EKDKQKQGHPL 287

RESULT 12

US-10-390-045-11
; Sequence 11, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POSTNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-390-045-11

Query Match 98.8%; Score 249; DB 14; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.2e-228;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSE 61
Db 1 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSE 60
Qy 62 STVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 121
Db 61 STVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 120
Qy 122 EPPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 181
Db 121 EPPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 180
Qy 182 TCYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 241
Db 181 TCYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 240
Qy 242 EKDKQKQGH 250
Db 241 EKDKQKQGH 249

RESULT 13

US-10-434-479-11
; Sequence 11, Application US/10434479
; Publication No. US20040092469A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; TITLE OF INVENTION: ANDROGEN-REGULATED PMP2A GENE AND POLYPEPTIDES
; FILE REFERENCE: 04995.0057-02000
; CURRENT APPLICATION NUMBER: US/10/434,479
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: 10/390,045
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-434-479-11

Query Match 98.8%; Score 249; DB 16; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.2e-228;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSE 61
Db 1 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRRREDALSSGCLWPSE 60
Qy 62 STVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 121
Db 61 STVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHFQPTPYLQHEIDLPTTISLSDGE 120
Qy 122 EPPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 181
Db 121 EPPYQGPCTQLRDPEQLELNRRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGISAT 180
Qy 182 TCYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 241
Db 181 TCYSGGGMGPPPTTSEVIGHYPCSSFOHQSSGPPSLLEGTRLHHTHIAPLESAATWS 240
Qy 242 EKDKQKQGH 250
Db 241 EKDKQKQGH 249

RESULT 14

US-10-000-256A-158
; Sequence 158, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes ar
; FILE REFERENCE: DEK-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapien

US-10-000-256A-158

Query Match 71.8%; Score 181; DB 14; Length 241;
Best Local Similarity 100.0%; Pred. No. 6.5e-164;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 72 PQVAPPRTDLAVPPFAQRERFQPTYPVLOHEIDLPTTISLSDGEEPPPYQGPCT 131
Db 60 PQVAPPRTDLAVPPFAQRERFQPTYPVLOHEIDLPTTISLSDGEEPPPYQGPCT 119
Qy 132 LQLRPEQOELNRESVRAPPNRTIFDSLDMSARLGCPSPSSNSGISATCVGSGRME 191
Db 120 LQLRPEQOELNRESVRAPPNRTIFDSLDMSARLGCPSPSSNSGISATCVGSGRME 179
Qy 192 GPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAIWSKROKQKQHP 251
Db 180 GPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAIWSKROKQKQHP 239
Qy 252 L 252
Db 240 L 240

RESULT 15

US-09-796-753-58
; Sequence 58, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19

; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 58
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Mouse
US-09-796-753-58
Query Match 20.2%; Score 51; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 3.6e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 111 LPPTISLSDGERPPPYQGPCTLQLRDPQOELNRESVRAPPNRTIFDSDL 161
Db 81 LPPTISLSDGERPPPYQGPCTLQLRDPQOELNRESVRAPPNRTIFDSDL 131
Search completed: May 20, 2004, 15:42:13
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:32:40 ; Search time 21 Seconds
(without alignments)
1154.298 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 252
Sequence: 1 MABLEFVQIIIVVMMVMV.....PLESAAIWSKEKDQKQKCHPL 252

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 78:*
1: Pirl1.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.2	308	2 A44496	transcription repr
2	8	3.2	582	2 S33937	hexon-associated p
3	8	3.2	794	2 T36972	probable membrane
4	8	3.2	1279	1 E69681	peptide synthetase
5	7	2.8	102	2 S04126	probable phospholi
6	7	2.8	110	4 S49808	hypothetical prote
7	7	2.8	116	2 A72685	hypothetical prote
8	7	2.8	161	2 B81245	conserved hypotet
9	7	2.8	220	2 A25513	25K silk protein -
10	7	2.8	267	2 T20096	hypothetical prote
11	7	2.8	271	2 J00950	ICP 18.5 protein -
12	7	2.8	279	2 G86406	unknown protein [i
13	7	2.8	289	2 I60170	syntaxin 3A - mous
14	7	2.8	296	2 T48926	bZIP transcription
15	7	2.8	297	2 T2980	conserved hypotet
16	7	2.8	297	2 H98302	hypothetical prote
17	7	2.8	334	2 B87545	hypothetical prote
18	7	2.8	339	2 S24161	lipase B - Pseudom
19	7	2.8	375	2 A75264	hypothetical prote
20	7	2.8	387	2 H72299	hypothetical prote
21	7	2.8	410	2 C96803	hypothetical prote
22	7	2.8	447	2 T16527	hypothetical prote
23	7	2.8	450	2 A7426	puc expression act
24	7	2.8	558	2 A23580	iron(III)-transpor
25	7	2.8	596	2 JC4697	trehalose trehaloh
26	7	2.8	603	2 T37518	probable transcrip
27	7	2.8	620	2 G90488	hypothetical prote
28	7	2.8	659	2 B84176	DNA mismatch repai
29	7	2.8	685	2 T09159	RNA helicase prh75

ALIGNMENTS

RESULT 1

A44496
transcription repressor factor AEF-1, adult enhancer factor 1 - fruit fly (Drosophila m.
C:Species: Drosophila melanogaster
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997
C:Accession: A44496
R:Falb, D.; Maniatis, T.
Mol. Cell. Biol. 12, 4093-4103, 1992
A>Title: Drosophila transcriptional repressor protein that binds specifically to negati
A:Reference number: A44496; MUID:92375078; PMID:1508206
A:Accession: A44496
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-308 <PAL>
A:Experimental source: ovary
A>Note: sequence extracted from NCBI backbone (NCBIN:111713, NCBIPI:111714)
C:Genetics:
A:Gene: FlyBase:Aef1
A:Cross-references: FlyBase:FBgn0005694
C:Keywords: DNA binding; transcription regulation

Query Match 3.2%; Score 8; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 212 QSSSGPPS 219

Db 75 QSSSGPPS 82

RESULT 2

S33937
hexon-associated protein - human adenovirus 12
C:Species: Mastadenovirus h12 (human adenovirus 12)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S33937
R:Sprenkel, J.
submitted to the EMBL Data Library, June 1993
A:Reference number: S33928
A:Accession: S33937
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <SPR>
A:Cross-references: EMBL:X73487; NID:g313361; PIDN:CAA51886.1; PID:g313371
C:Superfamily: adenovirus peripentonal hexon-associated protein

Query Match 3.2%; Score 8; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 AQRRERFRR 97

|||||

Db 134 AQRERFHR 141

RESULT 3

T36972

probable membrane associated protein - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36972
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21618
A:Accession: T36972
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-794 <OLI>
A:Cross-references: EMBL:AL109949; PIDN:CAB52886.1; GSPDB:GN00070; SCORDB:SCJ11.01c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB:SCJ11.01c

Query Match 3.2%; Score 8; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 ELNRESVR 149

Db 646 ELNRESVR 653

RESULT 4

E69681

peptide synthetase ppsE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Nov-2000
C:Accession: E69681
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, Iech, J.; Harwood, C.R.; Henaute, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Skiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; rognoni, A.; rosato, V.; Yoshiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69681
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1279 <KUN>
A:Cross-references: GB:Z99113; GB:AL009126; NID:G2634090; PIDN:CAB13713.1; PID:G2634213
A:Experimental source: strain 168
C:Genetics:
A:Gene: ppsE
C:Superfamily: peptide synthetase ppsE; acetate-CoA ligase homology; acyl carrier protein
C:Keywords: carrier protein; phosphotetrahine; phosphoprotein
E:522-960/Domain: acetate-CoA ligase homology <ACL>
E:978-1046/Domain: acyl carrier protein homology <ACP>
E:1010/Binding site: phosphotetrahine (Ser) (covalent) #status predicted

Query Match 3.2%; Score 8; DB 1; Length 1279;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 242 KEKDKQKG 249

Db 118 KEKDKQKG 125

RESULT 5

S04126

probable phospholipid transfer protein precursor, aleurone-specific - barley
N:Alternate names: 10K protein
C:Species: Hordeum vulgare (barley)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Sep-1997
C:Accession: S04126; S14610; S31457
R:Jakobsen, K.; Klemsdal, S.S.; Aalen, R.B.; Bosnes, M.; Alexander, D.; Olsen, O.A.
Plant Mol. Biol. 12, 285-293, 1989
A:Title: Barley aleurone cell development: molecular cloning of aleurone-specific cDNA
A:Reference number: S04126
A:Accession: S04126
A:Molecule type: mRNA
A:Residues: 1-102 <JAK>
A:Cross-references: EMBL:X15257; NID:G18892; PID:G18893
R:Kalla, R.; Klemsdal, S.S.; Linnestad, C.; Loenneberg, A.; Olsen, O.A.
submitted to the EMBL Data Library, January 1991
A:Description: A novel phospholipid transfer protein from immature barley aleurone cell
A:Reference number: S14610
A:Accession: S14610
A:Molecule type: mRNA
A:Residues: 1-102 <KAL>
A:Cross-references: EMBL:X57270; NID:G19042; PID:G19043
R:Kaller, R.; Potter, R.; Nielsen, P.S.; Linnestad, C.; Gabrielsen, O.S.; Olsen, O.A.
submitted to the EMBL Data Library, December 1992
A:Description: The promoter of the barley Ltp2 gene encoding an aleurone specific 7kd
A:Reference number: S31457
A:Accession: S31457
A:Molecule type: DNA
A:Residues: 1-102 <KAW>
A:Cross-references: EMBL:X69793; NID:G683753; PID:G19041
E:1-28/Domain: (or 31) signal sequence #status predicted <SIG>
E:29-102/Product: (or 32-102) probable phospholipid transfer protein, aleurone-specific.

Query Match 2.8%; Score 7; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VMMVMVV 21

Db 17 VMMVMVV 23

RESULT 6

S49808

hypothetical protein YN958.10 - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 05-Sep-1996 #text_change 20-Oct-2000
C:Accession: S49808
R:Devlin, K.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49800
A:Accession: S49808
A:Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-110 <DEV>
A:Cross-references: EMBL:Z46729; NID:G577134; PIDN:CAA86723.1; PID:G577144
C:Comment: There is no evidence that this sequence is expressed.
C:Genetics:
A:Map position: 13
C:Keywords: pseudogene

Query Match 2.8%; Score 7; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 IIIIVVV 15

Db 19 IIIIVVV 25

```

RESULT 7
A72685
hypothetical protein APE0897 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: A72685
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hailkawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Punahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-104, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72685
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <KAW>
A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79881.1; PID:d1043667; PID:G510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0897
C:Superfamily: Aeropyrum pernix hypothetical protein APE0897

Query Match 2.8%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 SARLGGP 170
Db 50 SARLGGP 56

RESULT 8
E81245
conserved hypothetical protein NMB0043 [imported] - Neisseria meningitidis (strain MC58)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
C:Accession: E81245; D82024
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizzi, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: E81245
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <TET>
A:Cross-references: GB:AE002364; NID:G7225269; PIDN:AAP40514.1; PID:G722526
A:Experimental source: serogroup B, strain MC58
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D82024
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <PAR>
A:Cross-references: GB:AL162752; GB:AL157959; NID:G7378778; PIDN:CAB83596.1; PID:G737905
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMB0043; NMA0289
C:Superfamily: Aquifex aeolicus conserved hypothetical protein aq_1996

Query Match 2.8%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 NRESVRA 150
Db 134 NRESVRA 140

```

```

RESULT 9
A25513
25K silk protein - silkworm
C:Species: Bombyx mori (silkworm)
C>Date: 30-Jun-1988 #sequence_revision 31-Dec-1988 #text_change 21-Jul-2000
C:Accession: A25513; A23020
R:Chevallard, M.; Couble, P.; Prudhomme, J.C.
Nucleic Acids Res. 14, 6341-6342, 1986
A>Title: Complete nucleotide sequence of the gene encoding the Bombyx mori silk protei
A:Reference number: A25513; MUID:86312902; PMID:3748814
A:Accession: A25513
A:Molecule type: DNA
A:Residues: 1-220 <CHE>
A:Cross-references: GB:X04226; GB:X02223; NID:G58988; PIDN:CAA27804.1; PID:G58989
R:Couble, P.; Chevallard, M.; Moine, A.; Ravel-Chapuis, P.; Prudhomme, J.C.
Nucleic Acids Res. 13, 1801-1814, 1985
A>Title: Structural organization of the P-25- gene of Bombyx mori and comparative anal
A:Reference number: A23020; MUID:85215581; PMID:2987837
A:Accession: A23020
A:Molecule type: protein
A:Residues: 1-41, '1', '43-81, 'Q', '83 <COU>
C:Genetics:
A:Introns: 18/1; 91/2; 137/1; 175/1
C:Superfamily: silkworm 25K silk protein

Query Match 2.8%; Score 7; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 ARSFISR 40
Db 205 ARSFISR 211

RESULT 10
T20096
hypothetical protein C50B8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T20096
R:Percy, C.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19223
A:Accession: T20096
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-267 <WIL>
A:Cross-references: EMBL:Z77654; PIDN:CAB01128.1; GSPDB:GN00023; CESP:C50B8.5
A:Experimental source: clone C50B8
C:Genetics:
A:Gene: CESP:C50B8.5
A:Map position: 5
A:Introns: 23/3; 118/3; 157/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C50B8.5

Query Match 2.8%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 PPPTYSE 199
Db 234 PPPTYSE 240

RESULT 11
JQ0950
ICP 18.5 protein - infectious laryngotracheitis virus (fragment)
C:Species: infectious laryngotracheitis virus
A>Note: host Gallus gallus (chicken)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Nov-2000
C:Accession: JQ0950
R:Griffin, A.M.
submitted to JIPID, January 1991

```

A;Reference number: JQ0950

A;Accession: JQ0950

A;Molecule type: DNA

A;Residues: 1-271 <GRI>

A;Experimental source: strain thorne V882

C;Genetics:

A;Gene: icp18.5

C;Superfamily: herpesvirus infected cell protein ICP18.5

Query Match 2.8%; Score 7; DB 2; Length 271;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 KLSARSP 37

|||||

DB 1 KLSARSP 7

RESULT 12

G86406

unknown protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: G86406

R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: G86406

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-279 <STO>

A;Cross-references: GB:AE005172; NID:gl1024879; PIDN:AAG26963.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match

Best Local Similarity 2.8%; Score 7; DB 2; Length 279;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 VITCLLS 28

|||||

DB 5 VITCLLS 11

RESULT 13

I60170

syntaxin 3A - mouse

C;Species: Mus musculus (house mouse)

C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 24-Sep-1999

C;Accession: I60170

R;Baraki, K.; Horikawa, H.P.; Morita, T.; Mori, H.; Sakimura, K.; Mishina, M.; Saisu, H.

Biochem. Biophys. Res. Commun. 211, 997-1005, 1995

A;Title: Identification of four different forms of syntaxin 3.

A;Reference number: I60170; MUID:95321968; PMID:7598732

A;Accession: I60170

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-289 <RES>

A;Cross-references: GB:D29797; NID:g924267; PIDN:EAA06180.1; PID:g924268

C;Superfamily: syntaxin

Query Match

Best Local Similarity 2.8%; Score 7; DB 2; Length 289;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IIIIVVV 15

|||||

DB 265 IIIIVVV 271

RESULT 14

T48926

bZIP transcription factor-like protein - Arabidopsis thaliana

N;Alternate names: protein F14L2.10

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C;Accession: T48926

R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseid, M.; Mewes, H.W.; Rudd, S.

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z25008

A;Accession: T48926

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-296 <JOR>

A;Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.10

A;Experimental source: cultivar Columbia; BAC clone F14L2

C;Genetics:

A;Gene: ATSP:F14L2.10

A;Map position: 3

A;Introns: 269/3

Query Match

Best Local Similarity 2.8%; Score 7; DB 2; Length 296;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PPSNSG 178

|||||

DB 125 PPSNSG 131

RESULT 15

AF2980

conserved hypothetical protein Atu3447 [imported] - Agrobacterium tumefaciens (strain

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AF2980

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl,

Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, R.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2980

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-297 <KUR>

A;Cross-references: GB:AE008689; PIDN:AL44260.1; PID:gl7741845; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3447

A;Map position: linear chromosome

Query Match

Best Local Similarity 2.8%; Score 7; DB 2; Length 297;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 RRREDAL 51

|||||

DB 34 RRREDAL 40

Search completed: May 20, 2004, 15:36:49

Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:50 ; Search time 18 Seconds
(without alignments)
728.982 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 252
Sequence: 1 MAELEFVQIIIVVMVMV.....PLESAAIWSKDKQKXGHEPL 252

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	99.2	287	1 TMEP HUMAN	Q969W9 homo sapien
2	51	20.2	260	1 TMEP MOUSE	Q9d7r2 mus musculus
3	35	13.9	306	1 C181 HUMAN	O15165 homo sapien
4	8	3.2	308	1 AEP1 DROME	P39413 drosophila
5	8	3.2	582	1 HEX3 ADEI2	P36712 human adeno
6	7	2.8	102	1 NUT2 HORVU	P20145 hordeum vul
7	7	2.8	145	1 CRAA ORYLA	P73919 oryzias lat
8	7	2.8	220	1 S125 BOMMO	P04148 bombyx mori
9	7	2.8	289	1 STK3 MOUSE	Q53238 rhizobium s
10	7	2.8	596	1 TRB2 RHISP	Q10086 schizosacch
11	7	2.8	603	1 YAO7 SCHPO	Q9ham6 halobacteri
12	7	2.8	708	1 MUTL HLAN1	Q9ugio homo sapien
13	7	2.8	778	1 TRAB HUMAN	Q8xk30 clostridium
14	7	2.8	778	1 GSH1 CLOPE	Q8qzq0 mus musculus
15	7	2.8	925	1 NPA3 MOUSE	Q8xf0 homo sapien
16	7	2.8	933	1 NPA3 MOUSE	P46012 caenorhabdi
17	7	2.8	952	1 YK15 CAEEL	P31629 homo sapien
18	7	2.8	1833	1 ZEP2 HUMAN	P44260 haemophilus
19	6	2.4	79	1 YF71 HAEIN	P33669 escherichia
20	6	2.4	86	1 YBBD ECOLI	Q9v5j8 homo sapien
21	6	2.4	90	1 IWL0 HUMAN	Q9nb33 aedes trise
22	6	2.4	95	1 GON1 PAGNA	P70074 pagrus major
23	6	2.4	103	1 RL44 AEDTR	P81730 achromobact
24	6	2.4	119	1 ACLY ACHLY	P59942 homo sapien
25	6	2.4	119	1 MCD1 HUMAN	Q58372 methanococc
26	6	2.4	128	1 Y962 METUA	Q9v3c7 homo sapien
27	6	2.4	131	1 SOH1 HUMAN	Q9xuhl schizosacch
28	6	2.4	131	1 SOH1 MOUSE	P18724 xenopus lae
29	6	2.4	139	1 SOH1 SCHPO	P03167 woodchuck h
30	6	2.4	140	1 ZG49 XENLA	P12914 woodchuck h
31	6	2.4	141	1 X WHV1	P12914 woodchuck h
32	6	2.4	141	1 X WHV59	P12914 woodchuck h
33	6	2.4	141	1 X WHVW6	P12914 woodchuck h

34	6	2.4	142	1 IBPB ECOLI	P29210 escherichia
35	6	2.4	142	1 MBAZ CLOAB	Q97h80 clostridium
36	6	2.4	144	1 NDK LYCES	P47921 lycopersico
37	6	2.4	148	1 SOH1 TAESO	Q8mpd6 taenia soli
38	6	2.4	148	1 YPX2 BLVJ	P03413 bovine leuk
39	6	2.4	149	1 Y4OM RHISN	P55598 rhizobium s
40	6	2.4	150	1 YP87 MYCPN	P75193 mycoplasma
41	6	2.4	151	1 TSR3 HUMAN	Q8i2w5 homo sapien
42	6	2.4	151	1 Y273 MERTH	Q26373 methanobact
43	6	2.4	155	1 MSRB XYLFA	Q9p229 xylella fas
44	6	2.4	155	1 MSRB XYLFA	Q87aj9 xylella fas
45	6	2.4	163	1 DEF3 SHEON	Q8ee60 shewanella

ALIGNMENTS

RESULT 1
TMEP_HUMAN
ID TMEP_HUMAN STANDARD; PRT; 287 AA.
AC Q969W9; Q96B72; Q9UDJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transmembrane prostate androgen-induced protein (Solid tumor-associated protein)
DE associated protein
GN TMEP1 OR PMPAL1 OR STAG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=20334621; PubMed=1087380;
RA Xu L.L., Shammugam N., Segawa T., Sesterhenn I.A., McLeod D.G., Moul J.W., Srivastava S.;
RA "A novel androgen-regulated gene, PMPAL, located on chromosome 20q13 exhibits high level expression in prostate."
RL Genomics 66:257-263 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=21453682; PubMed=11568975;
RA Rae P.K., Hooper J.D., Nicol D.L., Clements J.A.;
RT "Characterization of a novel gene, STAG1/PMPAL, upregulated in renal cell carcinoma and other solid tumors."
RL Mol. Carcinog. 32:44-53 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L., Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M., Clegg S., Cobley V.R., Collier R.R., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M., Ellington A.G., Griffiths J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths M.N.D., Griffiths M.N.D., Gwilliam R., Hall R.B., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leharvalho M.H., Levensha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Seward C.A., Sulston J.E., Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Trogans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

Qy	3	ELSFVQIIIVVMVVVVITCLSHYKLSARSFISRSQGRRRDLSSSECLWPSES	62
Dd	38	ELSFVQIIIVVMVVVVITCLSHYKLSARSFISRSQGRRRDLSSSECLWPSES	97
Qy	63	TVSNGIPEPQVYAPRPRTDRLAVPPPAQRPHRFPOPTYPYLQHEIDLPPTLSLDGEE	122
Dd	98	TVSNGIPEPQVYAPRPRTDRLAVPPPAQRPHRFPOPTYPYLQHEIDLPPTLSLDGEE	157
Qy	123	PPYPQGPCTQLQEDPQQOLELNESVRAPPNRTIFDSIDLWDSARIGSCPCPPSSNSGISAT	182
Dd	158	PPYPQGPCTQLQEDPQQOLELNESVRAPPNRTIFDSIDLWDSARIGSCPCPPSSNSGISAT	217
Qy	183	CYSGGRMECPPTTYSVIHYPGSSFGHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK	242
Dd	218	CYSGGRMECPPTTYSVIHYPGSSFGHQSSGPPSLLEGTRLHHTHIAPLESAAIWSK	277
Qy	243	EKDQKGHPGL 252	
Dd	278	EKDQKGHPGL 287	

RESULT 2

TMP MOUSE	ID	TMP MOUSE	STANDARD;	PRT;	260 AA.
AC	Q9D7R2; Q9EQH9;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Transmembrane prostate androgen-induced protein (Nedd4 WW domain-binding protein 4).				
GN	TMEPAI OR N4WBPA4.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OK	NCBI TaxID=10090;				
RX	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Stomach;				
KX	MEDLINE=21085660; PubMed=11217851;				
EA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Akakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyoosawa H., Kondo S., Yamanaoka I.,				
RA	Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadotani K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,				
RA	Hayaishizaki Y.;				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RE	Nature 409:685-690(2001).				
RN	[2]				
RP	SEQUENCE OF 59-260 FROM N.A.				
RX	MEDLINE=20498735; PubMed=11042109;				
RA	Jolliffe C.N., Harvey K.F., Haines B.P., Parasivam G., Kumar S.;				
RT	Identification of multiple proteins expressed in murine embryos as				
RT	binding partners for the WW domains of the ubiquitin-protein ligase				
RT	Nedd4.;				
RL	Biochem. J. 351:557-565(2000).				
CC	- - SUBUNIT: Interacts with the WW domains of NEDD4.				
CC	- - SUBCELLULAR LOCATION: Type Ib membrane protein (potential).				
CC	- - SIMILARITY: BELONGS TO THE TMEPAI FAMILY.				

This SWISS-PROT entry is copyright © It is produced through a collaboration between the EMBL, UniProt and the NCBI databases.

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AK008976; BAB26001.1; --
 DR EMBL; AF220208; AAG44247.1; --
 DR MGD; MGI:1929600; N4WBP4.
 DR GO; GO:0016021; C:integral to membrane; ISS.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0030521; P:androgen receptor signaling pathway; ISS.
 KW Transmembrane.
 FT DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 21 43 POTENTIAL.
 FT DOMAIN 44 260 CYTOPLASMIC (POTENTIAL).
 FT SITE 136 139 WW-BINDING (POTENTIAL).
 FT SITE 207 210 WW-BINDING (POTENTIAL).
 SQ SEQUENCE 260 AA; 28715 MW; 31AD07BD16B0D77D CRC64;

Query Match 20.2%; Score 51; DB 1; Length 260;
 Best Local Similarity 100.0%; Pred. No. 1.2e-45;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 LPPTLSLSDGEEPPPVQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSDL 161
 DB 124 LPPTLSLSDGEEPPPVQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSDL 174

RESULT 3

ID C181 HUMAN STANDARD; PRT; 306 AA.
 AC O15165; O15165-1; O15167; O15168;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein C18orf1.
 GN C18ORF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain;
 RX MEDLINE=98140124; PubMed=9479497;
 RA Yoshikawa T., Sanders A.R., Esterling L.E., Detera-Wadleigh S.D.;
 RT "Multiple transcriptional variants and RNA editing in C18orf1, a novel
 RT gene with LDLRA and transmembrane domains on 18p11.2";
 RL Genomics 47:246-257(1998).
 CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=Alpha-1;
 CC IsoId=O15165-1; Sequence=Displayed;
 CC Name=Alpha-2;
 CC IsoId=O15165-2; Sequence=VSP_006440;
 CC Name=Beta-1;
 CC IsoId=O15165-3; Sequence=VSP_006439;
 CC Name=Beta-2;
 CC IsoId=O15165-4; Sequence=VSP_006439, VSP_006440;
 CC -!- SIMILARITY: BELONGS TO THE TM6PAI FAMILY.
 CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF009424; AAC52023.1; --

DR EMBL; AF009425; AAC52024.1; --
 DR EMBL; AF009426; AAC52025.1; --
 DR EMBL; AF009427; AAC52026.1; --
 DR HSSP; P01130; IAJJ.
 DR Genew; HGNC:1224; C18orf1.
 DR MIM; 606571; --
 DR GO; GO:0016021; C:integral to membrane; NAS.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF000577; ldl_recept_a; 1.
 DR SMART; SM00192; LDLA; 1.
 DR PROSITE; PS01209; LDLRA_1; FALSE_NEG.
 DR PROSITE; PS00068; LDLRA_2; 1.
 KW Transmembrane; Alternative splicing.
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 85 POTENTIAL.
 FT DOMAIN 86 306 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 16 48 LDL-RECEPTOR CLASS A.
 FT VARSPLIC 1 61 MPEAGFOATNAPECTCTCTGKCLYGLSVLCVNQNDGDN
 SDEENCLAVTEHPHPGIPNS -> MAA (in isoform
 Beta-1 and isoform Beta-2).
 FT VARSPLIC 113 130 /FTId=VSP_006439.
 FT Missing (in isoform Alpha-2 and isoform
 Beta-2).
 FT /FTId=VSP_006440.
 SQ SEQUENCE 306 AA; 33900 MW; F48EF66E329201BD CRC64;

Query Match 13.9%; Score 35; DB 1; Length 306;
 Best Local Similarity 100.0%; Pred. No. 8.4e-29;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPVQGPCTQLQDRPEQQ 140
 DB 163 QHEIDLPTTISLSDGEEPPPVQGPCTQLQDRPEQQ 197

RESULT 4

ID AEP1 DROME STANDARD; PRT; 308 AA.
 AC P394I3; Q9VP37;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adult enhancer factor 1 (AEP-1).
 GN AEP1 OR CG5683.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Spheroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=92375078; PubMed=1508206;
 RA Falb D., Maniatis T.;
 RT "Drosophila transcriptional repressor protein that binds specifically
 RT to negative control elements in fat body enhancers";
 RL Mol. Cell. Biol. 12:4093-4103(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencosma P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulop D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson H., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celnik S.E.; cDNA resource.";
RT "A *Drosophila* full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR THAT BINDS SPECIFICALLY TO FAT
CC BODY-SPECIFIC ENHANCERS, NAMELY THE ADULT ADH ENHANCER (AAE) AND
CC THE ENHANCER THAT CONTROLS YOLK PROTEIN GENE EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: FOUND IN ALL TISSUES EXAMINED INCLUDING THE
CC OVARY AND THE FAT BODY.
CC -!- DEVELOPMENTAL STAGE: HIGHER LEVELS ARE FOUND IN THIRD-INSTAR
CC LARVAE AND IN ADULTS.
CC -!- SIMILARITY: Contains 4 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; M90755; -; NOT ANNOTATED_CDS.
DR EMBL; A8003594; AAF51722.1; -;
DR EMBL; AY060435; AAL25474.1; -;
DR PIR; A44496; A44496.
DR HSP; P08046; IAH.
DR TRANSFAC; T01513; -;
DR FlyBase; FBgn005694; Aef1.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
DR Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Zinc-finger; Metal-binding; Repeat.
FT DOMAIN 56 76 GLN-RICH (OPR-REPEAT).
FT DOMAIN 71 76 POLY-GLN.
FT DOMAIN 105 127 ALA-RICH.
FT DOMAIN 132 135 POLY-PRO.
FT ZN_FING 184 206 C2H2-TYPE 1.

FT ZN_FING 212 234 C2H2-TYPE 2.
FT ZN_FING 240 262 C2H2-TYPE 3.
FT ZN_FING 268 290 C2H2-TYPE 4.
SQ SEQUENCE 308 AA; 33782 MW; 54F04D90B6FC4250 CRC64;
Query Match 3.2%; Score 8; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 212 QSSSGPPS 219
DB 75 QSSSGPPS 82
RESULT 5
HEX3_ADE12 STANDARD; PRT; 582 AA.
ID HEX3_ADE12
AC F36712;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peripentonal hexon-associated protein (Protein IIIA).
GN PIIIA.
OS Human adenovirus type 12.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94076430; PubMed=8254750;
RA Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
RT "Nucleotide sequence of human adenovirus type 12 DNA: comparative
RT functional analysis.";
RL J. Virol. 68:379-389(1994).
RN [2]
RP SEQUENCE OF 566-576.
RX STRAIN=Huie;
RX MEDLINE=93174944; PubMed=8438575;
RA Freimuth P., Anderson C.W.;
RT "Human adenovirus serotype 12 virion precursors pMu and pVI are
RT cleaved at amino-terminal and carboxy-terminal sites that conform to
RT the adenovirus 2 endoprotease cleavage consensus sequence.";
RL Virology 193:348-355(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; X73487; CAA51886.1; -;
DR PIR; S33937; S33937.
DR InterPro; IPR003479; Hex_IIIa.
DR Pfam; PF02455; Hex_IIIa; 1.
KW Hexon-associated protein; Late protein.
SQ SEQUENCE 582 AA; 64786 MW; 8341875AE9A267F1 CRC64;
Query Match 3.2%; Score 8; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 90 AQERFHR 97
DB 134 AQERFHR 141
RESULT 6
NLT2_HORVU STANDARD; PRT; 102 AA.
ID NLT2_HORVU
AC P20145; 1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Probable nonspecific lipid-transfer protein precursor (LTP) (Aleurone-
 DE specific 10 kDa protein) (B-FABP).
 GN LTP2 OR B1E OR LTP.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticaceae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bomi;
 RA Jakobsen K., Klemmedal S.S., Aalen R.B., Bosnes M., Alexander D.,
 RA Olsen O.-A.;
 RA "Barley aleurone cell development: molecular cloning of aleurone-
 RT specific cDNAs from immature grains.";
 RL Plant Mol. Biol. 12:285-293(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bomi;
 RX MEDLINE=95152558; PubMed=7849757;
 RA Kalla R., Shimamoto K., Potter R., Nielsen P., Linnestad C.,
 RA Olsen O.-A.;
 RA "The promoter of the barley aleurone-specific gene encoding a
 RT putative 7 kDa lipid transfer protein confers aleurone cell-specific
 RT expression in transgenic rice.";
 RL Plant J. 6:849-860(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bomi; TISSUE=Leaf;
 RA Kalla R., Potter R., Nielsen P.S., Linnestad C., Gabrielsen O.S.,
 RA Olsen O.-A.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Potential phospholipid transfer protein.
 CC -!- TISSUE SPECIFICITY: ALEURONE.
 CC -!- DEVELOPMENTAL STAGE: MAXIMUM RNA ABUNDANCE AROUND MID-PHASE OF
 CC GRAIN DEVELOPMENT.
 CC -!- SIMILARITY: Belongs to the plant LTP family. B1E subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15257; CAA33329.1; -;
 DR EMBL; X57270; CAA40542.1; -;
 DR EMBL; X69793; CAA49448.1; -;
 DR FIR; S04126; S04126.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000528; Plant_LTP.
 DR Pfam; PF00234; tryp_alpha_aml; 1.
 DR SMART; SM00499; AAI; 1.
 DR PROSITE; PS00597; PLANT_LTP; FALSE NEG.
 KW Lipid-binding; Transport; Signal; Multigene family.
 FT SIGNAL 1 35
 FT CHAIN 36 102
 FT PROBABLE NONSPECIFIC LIPID-TRANSFER
 FT PROTEIN.
 FT DISULFID 37 71
 FT DISULFID 45 59
 FT DISULFID 60 95
 FT DISULFID 69 102
 FT BY SIMILARITY.
 SQ SEQUENCE 102 AA; 10357 MW; 474EEF35D0DD0558 CRC64;
 Query Match 2.8%; Score 7; DB 1; Length 102;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 VMVMVV 21
 DB 17 VMVMVV 23

RESULT 7
 CAA_ORYLA
 ID CAA_ORYLA STANDARD; PRT; 145 AA.
 AC 073919;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alpha crystallin A chain (Fragment).
 GN CRYAA.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Carolina Biological;
 RA Loosli F., Koester R.W., Carl M., Krone A., Wittbrodt J.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May contribute to the transparency and refractive index
 CC of the lens.
 CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ000940; CAA04397.1; -;
 DR InterPro; IPR001436; Crystallin_alpha.
 DR InterPro; IPR003090; Crystallin_N.
 DR InterPro; IPR002068; Hsp20.
 DR InterPro; IPR008978; HSP20_chap.
 DR Pfam; PF00525; crystallin; 1.
 DR PRINTS; PR00299; ACRYSTALLIN.
 DR ProDom; PD001193; Crystallin_N; 1.
 DR PROSITE; PS01031; HSP20; 1.
 KW Eye lens protein.
 FT NON_TER 1
 FT NON_TER 145
 FT SEQUENCE 145 AA; 16480 MW; 7B53CSADFD19706E CRC64;
 Query Match 2.8%; Score 7; DB 1; Length 145;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 SSNSGIS 180
 DB 39 SSNSGIS 45
 RESULT 8
 SI25_BOMMO
 ID SI25_BOMMO STANDARD; PRT; 220 AA.
 AC P04148;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 25 kDa silk protein precursor (p25).
 GN P25.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;

```

RN SEQUENCE FROM N.A.
RP STRAIN=703;
RX MEDLINE=86312902; PubMed=3749814;
RA Chevillard M., Couble P., Prudhomme J.-C.;
RT "Complete nucleotide sequence of the gene encoding the Bombyx mori
RT silk protein P25 and predicted amino acid sequence of the protein.";
RL Nucleic Acids Res. 14:6341-6342(1986).
[2]
RN SEQUENCE OF 1-83 FROM N.A.
RX MEDLINE=85215581; PubMed=2987837;
RA Couble P., Chevillard M., Moine A., Ravel-Chapuis P., Prudhomme J.-C.;
RT "Structural organization of the P25 gene of Bombyx mori and
RT comparative analysis of its 5' flanking DNA with that of the fibroin
RT gene.";
RL Nucleic Acids Res. 13:1801-1814(1985).
CC -1- SUBUNIT: Associates non-covalently with silk fibroin.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Produced exclusively in the posterior (PSG)
CC section of silk glands.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04226; CAA27804.1; -
CC PIR; A25513; A25513.
CC SIK; Glycoprotein; Signal.
CC CHAIN 1 17 BY SIMILARITY.
CC CARBOHYD 18 220 25 KDA SILK PROTEIN.
CC CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 220 AA; 25166 MW; 54D9C24C57D38ADA CRC64;
CC
CC Query Match 2.88; Score 7; DB 1; Length 220;
CC Best Local Similarity 100.0%; Pred. No. 15;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 34 ARSFISR 40
CC Db 205 ARSFISR 211
CC
CC RESULT 9
CC ID STX3 MOUSE STANDARD; PRT; 289 AA.
CC AC Q64704;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Syntaxin 3.
CC GN STX3A OR STX3.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC OX NCBI_TaxID=10090;
CC [1]
CC RN SEQUENCE FROM N.A. (ISOFORMS 3A; 3B; 3C; 3D; 3D-1 AND 3D-2).
CC RX STRAIN=ICR; TISSUE=Brain;
CC RX MEDLINE=95321968; PubMed=7598732;
CC RA Ibaraki K., Horikawa H.P.N., Morita T., Mori H., Sakimura K.,
CC Mishina K., Saito H., Abe T.;
CC RT "Identification of four different forms of syntaxin 3.";
CC RL Biochem. Biophys. Res. Commun. 211:997-1005(1995).
CC CC -1- FUNCTION: Potentially involved in docking of synaptic vesicles at
CC presynaptic active zones.
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:

```

```

CC Event=Alternative splicing; Named isoforms=6;
CC Name=3A;
CC IsoId=Q64704-1; Sequence=Displayed;
CC Name=3B;
CC IsoId=Q64704-2; Sequence=VSP_006345;
CC Name=3C;
CC IsoId=Q64704-3; Sequence=VSP_006341, VSP_006346;
CC Name=3D;
CC IsoId=Q64704-4; Sequence=VSP_006344;
CC Name=3D-1;
CC IsoId=Q64704-5; Sequence=VSP_006342;
CC Name=3D-2;
CC IsoId=Q64704-6; Sequence=VSP_006343;
CC -1- SIMILARITY: Belongs to the syntaxin/epimorphin family.
CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D29797; BAA06180.1; -
CC EMBL; D29798; BAA06181.1; ALT_TERM.
CC EMBL; D29799; BAA06182.1; -
CC EMBL; D29800; BAA06183.1; -
CC EMBL; D38375; BAA07454.1; ALT_SEQ.
CC PIR; I60170; I60170.
CC PIR; I83197; I83197.
CC PIR; I83198; I83198.
CC HSP; P32851; IBR0.
CC MGD; MGI.103077; Stx3.
CC InterPro; IPR006012; Syntaxin.
CC InterPro; IPR006011; Syntaxin_N.
CC InterPro; IPR000727; t-SNARE.
CC Pfam; PF05739; SNARE; 1.
CC Pfam; PF00804; Syntaxin; 1.
CC SMART; SM00503; SynN; 1.
CC SMART; SM00397; t-SNARE; 1.
CC PROSITE; PS00914; SYNTAXIN; 1.
CC PROSITE; PS50192; t-SNARE; 1.
CC Neurotransmitter transport; Coiled coil; Transmembrane;
CC Alternative splicing.
CC DOMAIN 1 263
CC TRANSMEM 264 283
CC
CC CYTOPLASMIC (POTENTIAL).
CC ANCHOR FOR TYPE IV MEMBRANE PROTEIN
CC (POTENTIAL).
CC EXTRACELLULAR (POTENTIAL).
CC COILED COIL (POTENTIAL).
CC T-SNARE COILED-COIL HOMOLGY.
CC IETRLNIDKISEHVEBAKLYSILSAPIPEPK -> NPH
CC GILSYLRLSSHE (in isoform 3C).
CC /FTid=VSP_006341.
CC TKDLEQLTTEIKK -> LPWNPLLSPEILT (in
CC isoform 3D-1).
CC /FTid=VSP_006342.
CC TKDLEQLTTEIKK -> LPRNPLLSPEILT (in
CC isoform 3D-2).
CC /FTid=VSP_006343.
CC Missing (in isoform 3D).
CC /FTid=VSP_006344.
CC EMLDNIELNVHTVHVHVKARDETRAMKYOGQARKLLII
CC IVVVVLLGILALIGLSVGLK -> AMIDRIENMDQSVG
CC FVERAVADTKKAVKYQSEARRKIMICCIILAILASTI
CC G (in isoform 3B).
CC /FTid=VSP_006345.
CC EMLDNIELNVHTVHVHVKARDETRAMKYOGQARKLLII
CC IVVVVLLGILALIGLSVGLK -> AMIDRIENMDQSVG
CC FVERAVADTKKAVKYQSEARRKIMICCIILAILASTI
CC GGIPA (in isoform 3C).
CC /FTid=VSP_006346.
CC SEQUENCE 289 AA; 33243 MW; 4A8EAPC3049EE6F CRC64;
CC SQ

```

Query Match 2.8%; Score 7; DB 1; Length 289;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 IIIIVV 15
 DB 265 IIIIVV 271

RESULT 10
 TREZ RHISP STANDARD; PRT; 596 AA.
 AC Q53238; (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Maltotriose trehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
 alpha-D-(1->4)-alpha-D-glucano)trehalose trehalohydrolase)
 DE (Maltotriose trehalose trehalohydrolase).
 GN TREZ.
 OS Rhizobium sp. (strain M-11).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI_TaxID=391;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96219094; PubMed=8829547;
 RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
 RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
 sp. M-11";
 RL Bioc. Biotechnol. Biochem. 60:717-720(1996).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
 linkage in 4-alpha-D-((1->4)-alpha-D-glucanoseyl)(n) trehalose to
 yield trehalose and alpha-(1->4)-D-glucan.
 CC -!- PATHWAY: Trehalose biosynthesis.
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; D78001; BAA11187.1; -
 PIR; JC4697; JC4697.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR Pfam; PF00128; alpha-amy1ase; 1.
 KW Hydrolase; Glycosidase.
 FT ACT SITE 265 265 BY SIMILARITY.
 SQ SEQUENCE 596 AA; 65262 MW; 73BA80AE0534DDCD CRC64;

Query Match 2.8%; Score 7; DB 1; Length 596;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 DALSESEG 55
 DB 287 DALSESEG 293

RESULT 11
 YAO7 SCHPO STANDARD; PRT; 603 AA.
 AC Q10086;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative transcriptional regulatory protein C11D3.07C.
 GN SPAC11D3.07C.
 OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle B.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
 RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerruti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; Z68166; CAA92308.2; -
 PIR; T37518; T37518.
 DR HSP; P08657; 1CLD.
 DR GeneDB SPombe; SPAC11D3.07C; -
 DR InterPro; IPR007219; Fungal_trans.
 DR InterPro; IPR001138; Fungi_Trn.
 DR Pfam; PF04082; Fungal_trans; 1.
 DR Pfam; PF00172; Zn clus; 1.
 DR PRINTS; PR00054; FUNGAL_ZNCS.
 DR SMART; SM00066; GAL4; 1.
 DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
 DR PROSITE; PS50048; ZN2_Cy6_FUNGAL_2; 1.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Zinc; Metal-binding.
 FT DNA_BIND 7 34 ZN(2)-CYS(6); FUNGAL-TYPE (POTENTIAL).
 SQ SEQUENCE 603 AA; 69815 MW; F4D08CA99B7FF415 CRC64;

Query Match 2.8%; Score 7; DB 1; Length 603;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 PTISLSD 119
 DB 421 PTISLSD 427

RESULT 12

```

MUTL HALNI
ID MUTL HALNI STANDARD; PRT; 659 AA.
AC Q9H5M6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE DNA mismatch repair protein mutL.
GN MUTL OR VNG0159G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahaling G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Ikenbarger T.A., Peck R.F., Pohlenschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is required for dam-dependent methyl-directed DNA mismatch
CC repair. May act as a "molecular matchmaker", a protein that
CC promotes the formation of a stable complex between two or more
CC DNA-binding proteins in an ATP-dependent manner without itself
CC being part of a final effector complex (By similarity).
CC -1- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A8004982; AAC18777.1; -
CC F1R; B84176; B84176.
CC HSSP; P23367; 1BKX.
CC HAMAP; MF 00149; -.
CC InterPro; IPR003594; AtPbind ATPase.
CC InterPro; IPR002099; DNA_mis_repair.
CC Pfam; PF01119; DNA_mis_repair; 1.
CC Pfam; PF02518; HATPase C; 1.
CC SMART; SM00387; HATPase C; 1.
CC TIGRFAMs; TIGR00585; mutL; 1.
CC PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
CC DNA repair; Complete proteome.
CC SEQUENCE 659 AA; 67917 MW; 16228484A85B9B8 CRC64;
Query Match 2.8%; Score 7; DB 1; Length 659;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 GRRRDA 50
DB 651 GRRRDA 657
|||||
RESULT 13
TRAB_HUMAN
ID TRAB_HUMAN STANDARD; PRT; 708 AA.
AC Q9UGIO;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE TRABID protein (Zinc finger RAN-binding domain containing 1).
GN ZRANB1 OR TRABID.
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL Medline=11792842; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RA "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + L-cysteine = ADP +
CC phosphate + gamma-L-glutamyl-L-cysteine.
CC -1- PATHWAY: Glutathione biosynthesis; first step.
CC -1- SIMILARITY: Belongs to the glutamate--cysteine ligase family 1.
CC
Query Match 2.8%; Score 7; DB 1; Length 708;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 116 SLSDGEE 122
DB 695 SLSDGEE 701
|||||
RESULT 14
GSH1_CLOPE
ID GSH1_CLOPE STANDARD; PRT; 778 AA.
AC Q8XK30;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glutamate--cysteine ligase (EC 6.3.2.2) (Gamma-
DE glutamylcysteine synthetase) (Gamma-BCS) (GCS).
DE GSHA OR CPE1573.
GN Clostridium perfringens.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RL STRAIN=13 / Type A;
RA MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RA "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + L-cysteine = ADP +
CC phosphate + gamma-L-glutamyl-L-cysteine.
CC -1- PATHWAY: Glutathione biosynthesis; first step.
CC -1- SIMILARITY: Belongs to the glutamate--cysteine ligase family 1.
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; AP003191; BAB81279.1; -;
HAWAP; MF_00578; -; 1.
DR InterPro; IPR007370; Glu cys ligase.
DR InterPro; IPR006335; Glu cys rel.
DR Pfam; PF04262; glu_cys_ligase_1.
DR TIGRFAMs; TIGR01435; glu_cys_lig_rel; 1.
KW Glutathione biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 778 AA; 90463 MW; 2008FF46AF06423E CRC64;

Query Match 2.8%; Score 7; DB 1; Length 778;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 115 ISLSDGE 121
| | | | |
DB 486 ISLSDGE 492

RESULT 15
ID_NPA3 MOUSE STANDARD; PRT; 925 AA.
AC Q90ZQ0; O9BQ04;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Neuronal PAS domain protein 3 (Neuronal PAS3) (Member of PAS protein
DE 6) (MOP6).
GN NPAS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20005801; PubMed=10534623;
RA Brunsell E.W., Witte D.P., Shreiner A.B., Potter S.S.;
RT "Characterization of Npas3, a novel basic helix-loop-helix PAS gene
RT expressed in the developing mouse nervous system.";
RL Mech. Dev. 88:237-241(1999).
RN [2]

PP SEQUENCE OF 141-280 FROM N.A.
RA Thomas R.S., Bradford C.A.;
RT "Cloning and chromosomal localization of MOP6.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May plays a broad role in neurogenesis.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Detected exclusively in adult brain.
CC -!- DEVELOPMENTAL STAGE: Expression detected between 9.5 and 11.5
CC dpc in the developing neural tube. Was also expressed
CC throughout the neuroepithelium of the developing central nervous
CC system between 10.5 and 12.5 dpc at 14.5 dpc. The
CC expression became restricted to the neopallial layer of the
CC cortex. At 12.5 dpc, expression was evident in nonneural
CC tissues such as the developing dermis and mesenchyme surrounding
CC the otic and nasal placodes. Expression was also detected in the
CC developing cardiac valves, limb and developing kidney.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; AF173871; AAF14283.1; -;
DR EMBL; AF168769; AAG35181.1; -;
DR MGI; MGI:1351610; Npas3.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00888; HLH; 1.
DR PROSITE; PS0112; PAS; 2.
KW Repeat; DNA-binding; Nuclear protein; Transcription regulation.
FT DNA_BIND 59 75
DOMAIN 76 112
PAS 1.
PAS 2.
FT DOMAIN 152 222
PAS 2.
FT DOMAIN 324 394
PAS 2.
FT DOMAIN 398 441
POLY-SER.
FT DOMAIN 236 244
POLY-GLY.
FT DOMAIN 706 709
POLY-GLY.
FT DOMAIN 760 771
R -> L (IN REF. 2).
FT CONFLICT 141 141
G -> E (IN REF. 2).
FT CONFLICT 156 156
F -> L (IN REF. 2).
FT CONFLICT 166 166
SQ SEQUENCE 925 AA; 100458 MW; B6B07C9B6755B938 CRC64;

Query Match 2.8%; Score 7; DB 1; Length 925;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 RMGGPPP 195
| | | | |
DB 124 RMGGPPP 130

Search completed: May 20, 2004, 15:35:17
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2004, 15:25:51 ; Search time 45 Seconds
(without alignments)
1766.902 Million cell updates/sec

Title: US-09-857-826b-17
Perfect score: 252
Sequence: 1 MABLEFVQIIIVVMMV.....PLESAIWSKDKQKHPL 252

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237	94.0	237	4 Q8NER4	Q8ner4 homo sapien
2	53	21.0	53	4 Q8BYL8	Q8byl8 homo sapien
3	51	20.2	84	4 Q8NTR9	Q8ntr9 homo sapien
4	35	13.9	200	11 Q8RSE2	Q8rse2 mus musculus
5	35	13.9	292	11 Q8BYE2	Q8bye2 mus musculus
6	35	13.9	306	11 Q8BWJ4	Q8bwj4 mus musculus
7	8	3.2	490	16 Q89B74	Q89b74 bradyrhizob
8	8	3.2	1117	16 Q8CK70	Q8ck70 streptomyce
9	8	3.2	1271	2 P94460	P94460 bacillus su
10	8	3.2	1279	16 Q31827	Q31827 bacillus su
11	7	2.8	51	16 Q8F917	Q8f917 leptospira
12	7	2.8	61	6 Q97510	Q97510 sus scrofa
13	7	2.8	73	12 Q8VB64	Q8vb64 white spot
14	7	2.8	95	4 Q86S99	Q86s99 homo sapien
15	7	2.8	96	10 Q8LRN2	Q8lrn2 ginkgo bilo
16	7	2.8	97	12 Q9P210	Q9p210 xestia c-ni

17	7	2.8	103	10 Q941S3	Q941s3 oryza sativ
18	7	2.8	116	17 Q9YDL6	Q9ydl6 acropyrum p
19	7	2.8	119	16 Q826M5	Q826m5 streptomyce
20	7	2.8	140	3 Q9P8K2	Q9p8k2 botrytis ci
21	7	2.8	140	11 Q35283	Q35283 rattus norv
22	7	2.8	145	2 Q34222	Q34222 vibrio chol
23	7	2.8	152	16 Q8DHA2	Q8dha2 synechococc
24	7	2.8	153	4 Q86US6	Q86us6 homo sapien
25	7	2.8	161	16 Q9URB6	Q9urb6 neisseria m
26	7	2.8	173	13 Q9DEV0	Q9dev0 clarias fus
27	7	2.8	176	9 Q9MBK1	Q9mbk1 bacterioph
28	7	2.8	179	16 Q8PF57	Q8pf57 xanthomonas
29	7	2.8	181	4 Q9NXP9	Q9npx9 homo sapien
30	7	2.8	182	13 Q7ZUF8	Q7zuf8 brachydanio
31	7	2.8	199	4 Q9Y4L8	Q9y4l8 homo sapien
32	7	2.8	203	17 Q8TQV5	Q8tqv5 methanosarc
33	7	2.8	208	5 Q18725	Q18725 caenorhabdi
34	7	2.8	220	5 Q9BL18	Q9bl18 bombyx mand
35	7	2.8	227	5 Q9GRX7	Q9grx7 calliphora
36	7	2.8	234	10 Q86E27	Q86e27 oryza sativ
37	7	2.8	239	17 Q976F2	Q976f2 sulfolobus
38	7	2.8	264	16 Q981E2	Q981e2 rhizobium l
39	7	2.8	271	12 Q69374	Q69374 gallid herp
40	7	2.8	278	12 Q91RB4	Q91rb4 papaya ring
41	7	2.8	278	12 Q8BD97	Q8bd97 papaya ring
42	7	2.8	278	12 Q8BD78	Q8bd78 papaya ring
43	7	2.8	278	12 Q8BD77	Q8bd77 papaya ring
44	7	2.8	278	12 Q8BD76	Q8bd76 papaya ring
45	7	2.8	278	12 Q8BD40	Q8bd40 papaya ring

ALIGNMENTS

RESULT 1

Q8NER4 PRELIMINARY; PRT; 237 AA.
AC Q8NER4; 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE FMEPAL variant A protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Brunschwig E.B., Wilson K., Mack D., Dawson D., Lawrence E.,
RA Willson J.K.V., Lu S., Nosrati A., Swinler S., Beard L.,
RA Lutterbaugh J.D., Willis J., Platzer P., Markowitz S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY128643; AAM89277.1; -
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0030521; P:androgen receptor signaling pathway; NAS.
SQ SEQUENCE 237 AA; 26201 MW; A4A274EABFD930 CRC64;

Query Match	94.0%;	Score 237;	DB 4;	Length 237;
Best Local Similarity	100.0%;	Pred. No. 4e-244;		
Matches 237;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	16	MMVVVVITCLSHYKLSARSFISRHSGQRRERDALSSGCLMPSESTVSGNGIPEQVY	75	
Db	1	MMVVVVITCLSHYKLSARSFISRHSGQRRERDALSSGCLMPSESTVSGNGIPEQVY	60	
Qy	76	APPPTDLRVLVPPAQRERHFPQTPYLOHEIDLPTTISLSDGEEPPVPGCTQLQR	135	
Db	61	APPPTDLRVLVPPAQRERHFPQTPYLOHEIDLPTTISLSDGEEPPVPGCTQLQR	120	
Qy	136	DPEQQLINRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGISATCYGSGRMEGPPP	195	
Db	121	DPEQQLINRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGISATCYGSGRMEGPPP	180	

QY 196 TVSEVIGHYGGSSFOHQSSGPPSLLEGTLRHHTHIAPIESAAIWSKSKOKQKHPL 252
|||||
Db 181 TVSEVIGHYGGSSFOHQSSGPPSLLEGTLRHHTHIAPIESAAIWSKSKOKQKHPL 237
|||||

RESULT 2

Q9BYL8 PRELIMINARY; PRT; 53 AA.
AC Q9BYL8; (FEB-2001) to the EMBL/GenBank/DBJ databases.
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE DJ1059L7.1.2 (Androgen induced type 1b transmembrane protein (PMEPAL),
isoform 2) (Fragment).
GN TMEPAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121913; CAC32857.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 6064 MW; A3B563F8B1F69782 CRC64;

Query Match 21.0%; Score 53; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.2e-48;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELFVQIIIVVMMVMMVVITCLSHYKLSARSFISRHSGRRRDALSS 53
|||||
Db 1 MAELFVQIIIVVMMVMMVVITCLSHYKLSARSFISRHSGRRRDALSS 53
|||||

RESULT 3

Q9NTR9 PRELIMINARY; PRT; 84 AA.
AC Q9NTR9; (FEB-2001) to the EMBL/GenBank/DBJ databases.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DJ1059L7.1.1 (Androgen induced type 1b transmembrane protein (PMEPAL),
isoform 1) (Fragment).
GN TMEPAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121913; CAB88144.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
FT NON_TER 84 84
SQ SEQUENCE 84 AA; 9236 MW; 7FCFCF1FFBFBFC92 CRC64;

Query Match 20.2%; Score 51; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.6e-46;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ELEFVQIIIVVMMVMMVVITCLSHYKLSARSFISRHSGRRRDALSS 53
|||||
Db 34 ELEFVQIIIVVMMVMMVVITCLSHYKLSARSFISRHSGRRRDALSS 84
|||||

RESULT 4

Q9R5E2 PRELIMINARY; PRT; 200 AA.
AC Q9R5E2; (FEB-2002) to the EMBL/GenBank/DBJ databases.
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Similar to chromosome 18 open reading frame 1 (Fragment).
GN D18ERTD653E OR D330030L18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022716; AAH22716.1; -.
DR MGD; MGI:1277150; D18ERTD653E.
FT NON_TER 1 1
SQ SEQUENCE 200 AA; 22117 MW; 14A78F6E8A4C09A5 CRC64;

Query Match 13.9%; Score 35; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQ 140
|||||
Db 57 QHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQ 91
|||||

RESULT 5

Q8BYE2 PRELIMINARY; PRT; 292 AA.
AC Q8BYE2; (FEB-2002) to the EMBL/GenBank/DBJ databases.
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protein C18ORF1 homolog.
GN D18ERTD653E OR D330030L18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK040289; BAC30562.1; -.
DR MGD; MGI:1277150; D18ERTD653E.
SQ SEQUENCE 292 AA; 32521 MW; F780D391984D66E3 CRC64;

Query Match 13.9%; Score 35; DB 11; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.6e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 QHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQ 140
|||||
Db 149 QHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQ 183
|||||

RESULT 6

Q8BWJ4 PRELIMINARY; PRT; 306 AA.
AC Q8BWJ4; (FEB-2002) to the EMBL/GenBank/DBJ databases.
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protein C18ORF1 homolog.
GN D18ERTD653E OR D330030L18RIK.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Heart;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK052338; BAC34943.1; --
 DR MGD; MGI:1277150; D18Ertd653e.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR SMART; SM00192; LDLa; 1.
 DR PROSITE; PS50068; LDLRA_2; 1.
 SQ SEQUENCE 306 AA; 33857 MW; 043C2260AEE8729F CRC64;
 Query Match 13.9%; Score 35; DB 11; Length 306;
 Best Local Similarity 100.0%; Pred. No. 1.7e-28;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 106 QHEIDLPTISLSDGHEPPYQGPCTQLQRDPQQ 140
 Db |||||||
 163 QHEIDLPTISLSDGHEPPYQGPCTQLQRDPQQ 197
 RESULT 7
 ID Q89EU9 PRELIMINARY; PRT; 490 AA.
 AC Q89EU9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE AprE protein.
 GN AprE BUK6971.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasamoto S., Watanabe A., Ideawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110.";
 RL DNA Res. 9:189-197(2002).
 DR EMBL; AP005960; BAC52236.1; --
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015428; F:type I protein secretor activity; IEA.
 DR GO; GO:0009306; P:protein secretion; IEA.
 DR InterPro; IPR006143; HlyVD.
 DR InterPro; IPR003997; RxdD.
 DR Pfam; PF00529; HlyD; 1.
 DR PRINTS; PR01490; RXTXKIND.
 KW Complete proteome.
 SQ SEQUENCE 490 AA; 54390 MW; AE3DF001D8F0F218 CRC64;
 Query Match 3.2%; Score 8; DB 16; Length 490;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 232 APLESAAI 239
 Db |||||||
 92 APLESAAI 99

RESULT 8
 ID Q8CK70 PRELIMINARY; PRT; 1117 AA.
 AC Q8CK70;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN SC00072 OR SCJ4.53C OR SCJ11.01C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939104; CAD55416.1; --
 DR PIR; T36972; T36972.
 DR InterPro; IPR006530; YD.
 DR Pfam; PF05593; RHS repeat; 9.
 DR TIGSFams; TIGR01643; YD repeat 2x; 9.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1117 AA; 119294 MW; C3AD8A5834EFB867 CRC64;
 Query Match 3.2%; Score 8; DB 16; Length 1117;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 142 ELNRESVR 149
 Db |||||||
 646 ELNRESVR 653
 RESULT 9
 ID P94460 PRELIMINARY; PRT; 1271 AA.
 AC P94460;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Peptide synthetase ORFS.
 GN PPS.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95227362; PubMed=7711903;
 RA Tognoni A., Franchi E., Magistrelli C., Colombo E., Cosmina P.,
 RA Grandi G.;
 RT "A putative new peptide synthase operon in Bacillus subtilis: partial
 RT characterization.";
 RL Microbiology 141:645-648(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Grandi G.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBSJ databases.

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA de Ferra F., Tognoni A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; 234883; CAA84364.1; -.
DR HSSP; P14687; IAMU.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016788; P:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR000873; AMP-binding.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006162; Ppantne S.
DR InterPro; IPR006163; Pp bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00668; Condensation; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00075; ACP DOMAIN; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR Phosphopantetheine.
SQ SEQUENCE 1271 AA; 144057 MW; 1F8B115E0B936C72 CRC64;

Query Match 3.2%; Score 8; DB 2; Length 1271;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 KKKQKQK 249
DB 111 KKKQKQK 118

RESULT 10
O31827 ID O31827 PRELIMINARY; PRT; 1279 AA.
AC O31827;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptide synthetase.
GN PPSE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst P., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Poulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

```

```

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandebol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst P., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 8-1279 FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048468; PubMed=9387222;
RA Tosato V., Albertini A.M., Zotti M., Sonda S., Bruschi C.V.;
RT "Sequence completion, identification and definition of the fengycin
RT operon in Bacillus subtilis 168."
RL Microbiology 143:3443-3450(1997).
DR EMBL; Z99113; CAB13713.1; -.
DR EMBL; Y13917; CAA74213.1; -.
DR HSSP; P14687; IAMU.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016788; P:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR000873; AMP-binding.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006162; Ppantne S.
DR InterPro; IPR006163; Pp bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00668; Condensation; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00075; ACP DOMAIN; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR Phosphopantetheine; Complete proteome.
SQ SEQUENCE 1279 AA; 144618 MW; 22EA638DEF0CEBE0 CRC64;

Query Match 3.2%; Score 8; DB 16; Length 1279;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 KKKQKQK 249
DB 118 KKKQKQK 125

RESULT 11
Q8F917 ID Q8F917 PRELIMINARY; PRT; 51 AA.
AC Q8F917;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA0208.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011210; AAN47407.1; -.
RW Hypothetical protein; Complete proteome.

```

SQ SEQUENCE 51 AA; 6014 MW; 27D5715PF73D1BAP CRC64;

Query Match 2.8%; Score 7; DB 16; Length 51;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AELSPVQ 8
Db 27 AELSPVQ 33

RESULT 12

O97510 PRELIMINARY; PRT; 61 AA.

AC O97510; (Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Syntaxin 3 homolog (Fragment).
OS Sus scrofa domestica (domestic pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9825;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUS=Thyroid;

RA Witke A., Koch M., Wahl R., Haring H.-U.;

RT "cDNA homologous to human Syntaxin 3.";

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF0041135; AAC79692.1; -

DR InterPro; IPR000727; T_SNARE.

DR Pfam; PF05739; SNARE; 1.

FT NON TER 1

SQ SEQUENCE 61 AA; 6865 MW; 4637D61BF80DF5AB CRC64;

Query Match 2.8%; Score 7; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 IIIIVV 15

Db 37 IIIIVV 43

RESULT 13

O8VB64 PRELIMINARY; PRT; 73 AA.

AC O8VB64; (Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Wsv124 (WSSV180).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
OX NCBI_TaxID=92652;
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21548311; PubMed=11689662;

RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

RT "Complete genome sequence of the shrimp white spot bacilliform

virus.";

RL J. Virol. 75:11811-11820(2001).

RN [2]

RP SEQUENCE FROM N.A.

RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Taiwan;

RX MEDLINE=20517549; PubMed=11062040;

RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.;

RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;

RT "Identification and characterization of a shrimp white spot syndrome

RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
cellular-type thymidine kinase and thymidylate kinase.";

RL Virology 277:100-110(2000).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=Taiwan;

RX MEDLINE=21844071; PubMed=11853398;

RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.;

RA Lo C.F.; Kou G.H.;

RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white

spot syndrome virus and characterization of the motif important for

targeting VP35 to the nuclei of transfected insect cells.";

RL Virology 293:44-53(2002).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=Taiwan;

RA Lo C.-F., Kou G.-H.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF332093; AAL33128.1; -

DR EMBL; AF440570; AAL39048.1; -

SQ SEQUENCE 73 AA; 8174 MW; AYD244DDIAC79C70 CRC64;

Query Match 2.8%; Score 7; DB 12; Length 73;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 171 CPPSSNS 177

Db 5 CPPSSNS 11

RESULT 14

Q86SR9

ID Q86SR9 PRELIMINARY; PRT; 95 AA.

AC Q86SR9;

DT 01-JUN-2003 (TRENBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE Similar to TRAP-binding protein domain.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUS=Colon, Kidney, and Stomach;

RA Strausberg R.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC048281; AAH48281.1; -

SQ SEQUENCE 95 AA; 11302 MW; 4C01331F90E1A45C CRC64;

Query Match 2.8%; Score 7; DB 4; Length 95;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 SLSGEE 122

Db 82 SLSGEE 88

RESULT 15

Q8LRN2

ID Q8LRN2 PRELIMINARY; PRT; 96 AA.

AC Q8LRN2;

DT 01-OCT-2002 (TRENBLrel. 22, Created)

DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Putative class II small heat shock protein (Fragment).

OS Ginkgo biloba (Ginkgo).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.

OX NCBI_TaxID=3311;

RN [1]

RP SEQUENCE FROM N.A.
 RA Ayadi R., Papon N., Chenieux J.-C., Rideau M.,
 RA Tremouillaux-Guiller J.,
 RT "Cloning, characterization and expression of a partial cDNA encoding a
 RT low molecular weight heat shock protein in Ginkgo biloba L.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF507963; AAW34241.1; -;
 DR GO; GO:0003773; F:heat shock protein activity; IEA.
 DR InterPro; IPR002068; Hsp20.
 DR InterPro; IPR008978; HSP20_chap.
 DR Pfam; PF00011; HSP20; 1.
 DR PROSITE; PS01031; HSP20; 1.
 KW Heat shock.
 FT NON_TER 1
 SQ SEQUENCE 96 AA; 10755 MW; 25B61BB503A03547 CRC64;
 Query Match 2.8%; Score 7; DB 10; Length 96;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 178 GISATCY 184
 Db 63 GISATCY 69

Search completed: May 20, 2004, 15:36:16
 Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 26, 2004, 14:55:02 ; Search time 6855 Seconds
(without alignments)
6708.528 Million cell updates/sec
Title: US-09-857-826b-44
Perfect score: 1061
Sequence: 1 tctcccttggttcgggtga.....cttcgcaaaaaaaaaaaaaa 1061
Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0
Searched: 3470272 seqs, 21671516995 residues
Word size : 0
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database :

Result No.	Score	Query Match	Length	DB ID	Description
1	1057	99.6	1085	6	AX775887 Sequence
2	961	90.6	969	6	BD272494 Secreted
3	955	90.0	1913	6	BD272544 Secreted
4	953	89.8	1061	9	BC015918 Homo sapi
5	953	89.8	1383	6	AX775889 Sequence
6	952	89.7	969	6	BD272514 Secreted
7	926	87.3	1141	9	AF224278 Homo sapi
8	912	86.0	1140	6	AR336830 Sequence
9	910	85.8	969	6	BD272515 Secreted
10	910	85.8	969	6	BD272516 Secreted
11	903	85.1	1818	9	AY128643 Homo sapi
12	902	85.0	4839	9	AF305616 Homo sapi
13	756	71.3	756	6	BD272495 Secreted
14	753	71.0	753	6	BD272534 Secreted
15	749	70.6	1321	6	AX392417 Sequence
16	708	66.7	759	6	BD272545 Secreted
17	708	66.7	759	6	AR336831 Sequence
18	701	66.1	861	6	AX392419 Sequence
19	695	65.5	61505	9	AF305426 Homo sapi
20	695	65.5	130435	9	HS718J7 Human DNA
21	509	48.0	1583	6	AX593655 Sequence
22	350	33.0	408	6	AX071267 Sequence
23	322	30.3	812	6	AX011709 Sequence
24	322	30.3	812	6	BD226320 Pancrati
25	313	29.5	693	6	AX392430 Sequence
26	157	14.8	150224	9	HSJ1059L7 Human DNA
27	79	7.4	2570	9	AK056098 Homo sapi
28	51	4.8	51	6	AX199565 Sequence
29	44	4.1	411	6	BD272547 Secreted
30	44	4.1	484	6	BD272546 Secreted
31	44	4.1	648	6	BD272535 Secreted
32	44	4.1	651	6	BD272505 Secreted
33	44	4.1	651	10	AF220208 Mus muscu
34	44	4.1	878	6	AX392428 Sequence
35	44	4.1	1379	10	BC036995 Mus muscu
36	44	4.1	1713	6	BD272504 Secreted
37	44	4.1	1713	6	BD272517 Secreted
38	44	4.1	1713	6	BD272518 Secreted
39	44	4.1	1713	6	BD272519 Secreted
40	44	4.1	156698	10	AL837509 Mouse DNA
41	44	4.1	175754	2	AC110189 Mus muscu
42	44	4.1	231930	2	AC134911 Mus muscu
43	42	4.0	673	6	AX525744 Sequence
44	42	4.0	249554	2	AC139417 Rattus no
45	42	4.0	258632	2	AC111878 Rattus no

ALIGNMENTS

RESULT 1
AX775887
LOCUS AX775887
DEFINITION Sequence 157 from Patent WO03048202.
ACCESSION AX775887
VERSION AX775887.1 GI:32693605
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 157 12-JUN-2003;
Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Qy	781	CCA	CAC	ACATCGCGCCCTTAGAGCGCAGCCATCTCGAGCAAAAGAGAGATTAACA	840
Db	808	CCA	CAC	ACATCGCGCCCTTAGAGCGCAGCCATCTCGAGCAAAAGAGAGATTAACA	867
Qy	841	GA	AAAGACACCCCTCTCTAGGGTCCCACAGGGGGGCGCGGCTGGGGCTCGTAGCTGA	900	
Db	868	GA	AAAGACACCCCTCTCTAGGGTCCCACAGGGGGGCGCGGCTGGGGCTCGTAGCTGA	927	
Qy	901	GC	AGAACATCTCGCGCTTCTTAGAGAGGAGTGTAGAGGGAAGCGCGGGGGCGCAGCAACGC	960	
Db	928	GC	AGAACATCTCGCGCTTCTTAGAGAGGAGTGTAGAGGGAAGCGCGGGGGCGCAGCAACGC	987	
Qy	961	AT	CGTGTGGCGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGATCTCTGGTC	1020	
Db	988	AT	CGTGTGGCGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGATCTCTGGTC	1047	
Qy	1021	TGA	ATGCACAAGCTTAAGAGAGCTTGC	1057	
Db	1048	TGA	ATGCACAAGCTTAAGAGAGCTTGC	1084	

Query Match 99.6%; Score 1057; DB 6; Length 1085;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1057; Conservative 0; Mismatches 0;
Gaps 0;

RESULT 2	BD272494	969 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272494	Secreted proteins and nucleic acids encoding them.			
DEFINITION	BD272494				
ACCESSION	BD272494				
VERSION	BD272494.1	GI:33082262			
KEYWORDS	JP 2002539773-A/3.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 969)				
JOURNAL	Barnes,T.M., Holtzman,D.A., Sharp,J.D. and Fraser,C.C.				
	Secreted proteins and nucleic acids encoding them				
	Patent: JP 2002539773-A 3 26-NOV-2002;				
	MILLENNIUM PHARMACEUTICALS INC				
COMMENT	OS Homo sapiens (human)				
	PN JP 2002539773-A/3				
	PD 26-NOV-2002				
	PF 01-MAR-2000 JP 2000602247				
	PR 01-MAR-1999 US 60/122458				
	PI THOMAS M BARNES,DOUGLAS A HOLTZMAN,JOHN D SHARP,CHRISTOPHER C				
	PI FRASER				
	PC C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/68,PC				
	G01N33/15,				
	PC G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC				
	Secreted proteins and nucleic acids encoding them FH Key				
	Location/Qualifiers				
FT	CDS	(6)..(761).			
FEATURES	source	Location/Qualifiers			
		1..969			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
ORIGIN					
	Query Match	90.6%; Score 961; DB 6; Length 969;			
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches 961; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Qy	101	ATGCGGAGCTGAGTTTGTTCAGATCATCATCATCTGGTGGTGCATGATGGTGCATGGTGC	160		
Db	6	ATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCTGGTGGTGCATGATGGTGCATGGTGC	65		
Qy	161	GTGGTGATCACGTGCTGTGAGCCACTACAAGCTGTGTGCACGGTCTTTCATCAGCCGG	220		
Db	66	GTGGTGATCACGTGCTGTGAGCCACTACAAGCTGTGTGCACGGTCTTTCATCAGCCGG	125		
Qy	221	CACGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGATGCCCTGTGGCCCTCG	280		
Db	126	CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGATGCCCTGTGGCCCTCG	185		

QY 281 GAGAGCAGAGTGTGAGGCAACGGAATCCAGAGCCGCGAGTCTACCGCCCGCTCGGCC 340
Db 186 GAGAGCAGAGTGTGAGGCAACGGAATCCAGAGCCGCGAGTCTACCGCCCGCTCGGCC 245
QY 341 ACCGACCGCTGGCGGTGCGCCCTTCCGCGAGGAGCGCTTCCACCGCTTCCAGGCC 400
Db 246 ACCGACCGCTGGCGGTGCGCCCTTCCGCGAGGAGCGCTTCCACCGCTTCCAGGCC 305
QY 401 ACCTATCGTACCTGACGACGAGATCGACTGCGCCGCCACCATCTCGCTGTGACAGCGG 460
Db 306 ACCTATCGTACCTGACGACGAGATCGACTGCGCCGCCACCATCTCGCTGTGACAGCGG 365
QY 461 GAGAGCCCGCCACCTTACAGGCGCCCTGACCTTCCAGCTTCCGAGCCCGAGCAGAG 520
Db 366 GAGAGCCCGCCACCTTACAGGCGCCCTGACCTTCCAGCTTCCGAGCCCGAGCAGAG 425
QY 521 CTGGAATCTGAACCGGAGTGGTGGCGGCGACCCCGCAACAGAACCATCTTCCAGAGTGC 580
Db 426 CTGGAATCTGAACCGGAGTGGTGGCGGCGACCCCGCAACAGAACCATCTTCCAGAGTGC 485
QY 581 CTGATGATAGTGTGAGGCGCCCTGCGCCCGCCCGCAGAGTAACCTCGGCGATCAGC 640
Db 486 CTGATGATAGTGTGAGGCGCCCTGCGCCCGCCCGCAGAGTAACCTCGGCGATCAGC 545
QY 641 GCCAGTGTACGAGCGCGCGCGATGAGAGGCGCGCGCCCGCAGAGTGTGACAGAGTGC 700
Db 546 GCCAGTGTACGAGCGCGCGCGATGAGAGGCGCGCGCCCGCAGAGTGTGACAGAGTGC 605
QY 701 ATCGGCCACTACCGGCGTCTCTTCCAGACACAGAGAGAGTGGGCGCGCTCTCTTG 760
Db 606 ATCGGCCACTACCGGCGTCTCTTCCAGACACAGAGAGAGTGGGCGCGCTCTCTTG 665
QY 761 CTGAGGCGGACCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 820
Db 666 CTGAGGCGGACCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 725
QY 821 AGCAAGAGAGGATTAACAGAAAGGAGCACTCTCTAGGCTCCCGAGGCGCGCGCT 880
Db 726 AGCAAGAGAGGATTAACAGAAAGGAGCACTCTCTAGGCTCCCGAGGCGCGCGCT 785
QY 881 GGGCTGGTGTGAAAGGAGCACTCTCTAGGCTCCCGAGGAGAGTGTGAGAGGAA 940
Db 786 GGGCTGGTGTGAAAGGAGCACTCTCTAGGCTCCCGAGGAGAGTGTGAGAGGAA 845
QY 941 GGGCGGGGCGCAGCAACGATGTGTGGCGCTTCCCTCCACCTCTGTGTATTAATA 1000
Db 846 GGGCGGGGCGCAGCAACGATGTGTGGCGCTTCCCTCCACCTCTGTGTATTAATA 905
QY 1001 TTTATATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060
Db 906 TTTATATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 965
QY 1061 A 1061
Db 966 A 966

RESULT 3
BD272544
LOCUS 1913 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and nucleic acids encoding them.
ACCESSION BD272544
VERSION BD272544.1 GI:33082312
KEYWORDS JP 2002539773-A/53.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1913)
Barnes,T.M., Holtzman,D.A., Sharp,J.D. and Fraser,C.C.
Secreted proteins and nucleic acids encoding them
Patent: JP 2002539773-A 53 26-NOV-2002;
JOURNAL

MILLENNIUM PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002539773-A/53
PD 26-NOV-2002
PF 01-MAR-2000 JP 2000602247
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/69, PC
G01N33/15,
G01N33/50, G01N33/53, G01N33/56, C12N15/00, C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH Key
Location/Qualifiers
FT source 1..1913
FT Location/Qualifiers
1..1913
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
ORIGIN
Query Match 90.0%; Score 955; DB 6; Length 1913;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1055; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TCCTCTTGGGTTGGGTTGAAAGCGCTTGGGGTTGAGTGGGCCATGATCCCGAGCTGC 60
Db 46 TCCTCTTGGGTTGGGTTGAAAGCGCTTGGGGTTGAGTGGGCCATGATCCCGAGCTGC 105
QY 61 TGGGAATCTGAGGCGGAGCGCTCTCTCGGAAACAGGCAATGGCGGAGCTGGATTGCT 120
Db 106 TGGGAATCTGAGGCGGAGCGCTCTCTCGGAAACAGGCAATGGCGGAGCTGGATTGCT 165
QY 121 TCAGATCATCATCATCTGCTGGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 166 TCAGATCATCATCATCTGCTGGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 225
QY 181 GAGCCTATCAAGCTGTCTGCAAGTCTTCTTATCATAGCGGCAACAGCGGCGGAGGAG 240
Db 226 GAGCCTATCAAGCTGTCTGCAAGTCTTCTTATCATAGCGGCAACAGCGGCGGAGGAG 285
QY 241 AGAAGATCCCTGCTCTCAGAGAGTGTCTGTGGCCCTCGGAGAGCAGAGTGTCAAGCAA 300
Db 286 AGAAGATCCCTGCTCTCAGAGAGTGTCTGTGGCCCTCGGAGAGCAGAGTGTCAAGCAA 345
QY 301 CGGAATCCAGAGCGCAGGTCTAGCGCCCGCTCGGCGCCACCGAGCGCTGGCCGTGCC 360
Db 346 CGGAATCCAGAGCGCAGGTCTAGCGCCCGCTCGGCGCCACCGAGCGCTGGCCGTGCC 405
QY 361 GCGCTTGGCCAGCGGAGCGCTTTCACCGCTTCCAGCGCCACCTATCGGTACCTGAGCA 420
Db 406 GCGCTTGGCCAGCGGAGCGCTTTCACCGCTTCCAGCGCCACCTATCGGTACCTGAGCA 465
QY 421 CGAGATCGACTGTGCGGCGCCACCATCTCGCTGTGAGCGGAGGAGCGCCCGCTTACCA 480
Db 466 CGAGATCGACTGTGCGGCGCCACCATCTCGCTGTGAGCGGAGGAGCGCCCGCTTACCA 525
QY 481 GGGCGCCCTGCACTTCCAGCTTGGGAGCCCGAGCAGCAGCTGTGAACTGAAACCGGAGTGC 540
Db 526 GGGCGCCCTGCACTTCCAGCTTGGGAGCCCGAGCAGCAGCTGTGAACTGAAACCGGAGTGC 585
QY 541 GGTGCGGCAACCCCAACAGAACCATCTTTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 600
Db 586 GGTGCGGCAACCCCAACAGAACCATCTTTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 645
QY 601 GGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 646 GGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
QY 661 GGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 706 GGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 765

|||||
 330 ATCCGTAATGCGAGCGAGATGAGCTGCTCCGCCACCACTCTGCTGTGAGCGGGAGG 449
 |||||
 465 AGCCCCACCCCTACAGGCGCCCTCCAGCTTCAGCTTCGGGACCCCGAGCAGAGCTGG 524
 |||||
 450 AGCCCCACCCCTACAGGCGCCCTCCAGCTTCAGCTTCGGGACCCCGAGCAGAGCTGG 509
 |||||
 525 AACTGAACCGGGAGTGGGTCGGGCGACCCCGAAGACAGACCACTTCGACAGTGA 584
 |||||
 510 AACTGAACCGGGAGTGGGTCGGGCGACCCCGAAGACAGACCACTTCGACAGTGA 569
 |||||
 585 TGGATAGTGGCAGGCTGGGCGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATTACGCGCA 644
 |||||
 570 TGGATAGTGGCAGGCTGGGCGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATTACGCGCA 629
 |||||
 645 CTGTCTACGCGAGCGCGGGCGGCATGAGAGGGCGCGCGCCACCTACAGCGAGGTCATCG 704
 |||||
 630 CTGTCTACGCGAGCGCGGGCGGCATGAGAGGGCGCGCGCCACCTACAGCGAGGTCATCG 689
 |||||
 705 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGCAGTGGGCGCGCCCTCTTCTGCTGG 764
 |||||
 690 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGCAGTGGGCGCGCCCTCTTCTGCTGG 749
 |||||
 765 AGGGGACCCGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCATTCTGGAGCA 824
 |||||
 750 AGGGGACCCGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCATTCTGGAGCA 809
 |||||
 825 AAGAGAAGGATAAACAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 884
 |||||
 810 AAGAGAAGGATAAACAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 869
 |||||
 885 CTCGTAGGTGAAGAGCAGACACTCGCGGCTCTTAGAGAGAGAGTGAAGAGAGGGG 944
 |||||
 870 CTCGTAGGTGAAGAGCAGACACTCGCGGCTCTTAGAGAGAGAGTGAAGAGAGGGG 929
 |||||
 945 GGGGCGCAGCAACGATCTGTGGGCTCCCTCCACCTCCCTCTGTGTATATAATTTA 1004
 |||||
 930 GGGGCGCAGCAACGATCTGTGGGCTCCCTCCACCTCCCTCTGTGTATATAATTTA 989
 |||||
 1005 CATGTGATGTCTGGTCTGAATGACACAGCTAAGAGAGCTTCGCAAAAAA 1057
 |||||
 990 CATGTGATGTCTGGTCTGAATGACACAGCTAAGAGAGCTTCGCAAAAAA 1042
 |||||

RESULT 5
 AX775889
 LOCUS 1383 bp mRNA linear PAT 14-JUL-2003
 DEFINITION Sequence 159 from Patent WO03048202.
 ACCESSION AX775889
 VERSION AX775889.1 GI:32693607
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Matsuda, A. and Muramatsu, S.
 NP-kB activating gene
 Patent: WO 03048202-A 159 12-JUN-2003;
 Asahi Kasei Kabushiki Kaisha (JP)
 FEATURES
 Location/Qualifiers
 1..1383
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 321..1184
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAB11644.1"
 /db_xref="GI:32693608"
 /translation="MHRLMGVNSTAAAAAGQPNVSCNCKRSIFQSMETELBFVQI
 IIVVMVMVWVITCLLSHYKLSARSPISRSQRRREDALSGELWPSESTVSGN
 GIPEPVVAPRPRTDLAVPPAQRERPHRPQTYPLQHBIDLPPPTISLSDGEPPPP
 source
 CDS

YQCPCTQLRDPQOQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSPNSGISTATC
 YGSGRMEGPPPTYSYSEVIGHYFSSGFHQHSSGPPSLBCTRLHHTHIAPLESNAIWS
 KKKDKQKQHPL*
 Query Match 89.8%; Score 953; DB 6; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 105 CGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGGTGGTGG 164
 DB 430 CGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGGTGGTGG 489
 QY 165 TGATCATCTGCTGCTGAGCCACTACAGCTCTCTGCAAGTCTCTTCATCAGCGGACCA 224
 DB 490 TGATCATCTGCTGCTGAGCCACTACAGCTCTCTGCAAGTCTCTTCATCAGCGGACCA 549
 QY 225 GCCAGGGCGGAGGAGAGAGATGCCCTGCTCTCAGAGGATGCTGTGGCCCTTCGAGAGA 284
 DB 550 GCCAGGGCGGAGGAGAGAGATGCCCTGCTCTCAGAGGATGCTGTGGCCCTTCGAGAGA 609
 QY 285 GCACTAGTGTGAGGCAACCGGAATCCAGAGCCCGAGGTCTAAGCCCGCCCTTCGCGCCCA 344
 DB 610 GCACTAGTGTGAGGCAACCGGAATCCAGAGCCCGAGGTCTAAGCCCGCCCTTCGCGCCCA 669
 QY 345 ACCGCTTGGCGCTGCGCCCTTCGCGCCAGCGGGAGCGCTTCACCGCTTCAGCGCCACCT 404
 DB 670 ACCGCTTGGCGCTGCGCCCTTCGCGCCAGCGGGAGCGCTTCACCGCTTCAGCGCCACCT 729
 QY 405 ATCCCTTACCTGAGCAGCAGAGATCGGACCTGCGCGCCACCATCTGCTGTGTGAGAGG 464
 DB 730 ATCCCTTACCTGAGCAGCAGAGATCGGACCTGCGCGCCACCATCTGCTGTGTGAGAGG 789
 QY 465 AGCCCTTACCTGAGCAGCAGAGATCGGACCTGCGCGCCACCATCTGCTGTGTGAGAGG 524
 DB 750 AGCCCTTACCTGAGCAGCAGAGATCGGACCTGCGCGCCACCATCTGCTGTGTGAGAGG 849
 QY 525 AACTGAACCGGGAGTGGGTCGGGCGACCCCGAAGACAGACCACTTCGACAGTGA 584
 DB 850 AACTGAACCGGGAGTGGGTCGGGCGACCCCGAAGACAGACCACTTCGACAGTGA 909
 QY 585 TGGATAGTGGCAGGCTGGGCGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATTACGCGCA 644
 DB 910 TGGATAGTGGCAGGCTGGGCGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATTACGCGCA 969
 QY 645 CTGTCTACGCGAGCGCGGGCGGCATGAGAGGGCGCGCGCCACCTACAGCGAGGTCATCG 704
 DB 970 CTGTCTACGCGAGCGCGGGCGGCATGAGAGGGCGCGCGCCACCTACAGCGAGGTCATCG 1029
 QY 705 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGCAGTGGGCGCGCCCTCTTCTGCTGG 764
 DB 1030 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGCAGTGGGCGCGCCCTCTTCTGCTGG 1089
 QY 765 AGGGGACCCGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATTCTGGAGCA 824
 DB 1090 AGGGGACCCGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATTCTGGAGCA 1149
 QY 825 AAGAGAAGGATAAACAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 884
 DB 1150 AAGAGAAGGATAAACAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 1209
 QY 885 CTGTCTAGGTGAAGAGCAGACACTCGCGGCTCTTAGAGAGAGAGTGAAGAGAGGGG 944
 DB 1210 CTGTCTAGGTGAAGAGCAGACACTCGCGGCTCTTAGAGAGAGAGTGAAGAGAGGGG 1269
 QY 945 GGGGCGCAGCAACGATCTGTGGGCTCCCTCCACCTCCCTCTGTGTATATAATTTA 1004
 DB 1270 GGGGCGCAGCAACGATCTGTGGGCTCCCTCCACCTCCCTCTGTGTATATAATTTA 1329
 QY 1005 CATGTGATGTCTGGTCTGAATGACACAGCTAAGAGAGCTTCGCAAAAAA 1057
 DB 1330 CATGTGATGTCTGGTCTGAATGACACAGCTAAGAGAGCTTCGCAAAAAA 1382

QY	590	AGTGCAGGCTGGGGCGGCGCCCTGCGCCCGCCAGCAGTAAGTCTGGGGCATCAGCGCCACGTCG	649
Db	495	AGTGCAGGCTGGGGCGGCGCCCTGCGCCCGCCAGCAGTAAGTCTGGGGCATCAGCGCCACGTCG	554
QY	650	TACGGCAGCGGGCGGGCGCATGAGAGGGGCGCGCCACCTACAGCGAGGTTCATCGGCCAC	709
Db	555	TACGGCAGCGGGCGGGCGCATGAGAGGGGCGCGCCACCTACAGCGAGGTTCATCGGCCAC	614
QY	710	TACCGGGGGTCTGCTTCACGACCCAGCAGGAGCAGTGGGGCGCGCCCTCTCTCTCTGAGAGGG	769
Db	615	TACCGGGGGTCTGCTTCACGACCCAGCAGGAGCAGTGGGGCGCGCCCTCTCTCTCTGAGAGGG	674
QY	770	ACCCGGCTCCACACACACATCGCGCCCTCTAGAGAGCGCAGCCATCTGGAGCAAAAGAG	829
Db	675	ACCCGGCTCCACACACACATCGCGCCCTCTAGAGAGCGCAGCCATCTGGAGCAAAAGAG	734
QY	830	AAGGATAACAGAAAGAGCACCTCTTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGG	889
Db	735	AAGGATAACAGAAAGAGCACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGG	794
QY	890	TAGGTGAAAGAGCAGAACACCTCGCGCTCTCTTAGAAGAGGAGTCTAGAGAGGCGGGGG	949
Db	795	TAGGTGAAAGAGCAGAACACCTCGCGCTCTCTTAGAAGAGGAGTCTAGAGAGGCGGGGG	854
QY	950	CGCAGCAACGATCGTGTGGCGCCCTCCCTCCACCTCCCTCTGTGTATAAATTTTTCATGT	1009
Db	855	CGCAGCAACGATCGTGTGGCGCCCTCCCTCCACCTCCCTCTGTGTATAAATTTTTCATGT	914
QY	1010	GATGTCGTGTCGAATGACAAAGCTAGAGAGCTTGCAAAAAA	1061
Db	915	GATGTCGTGTCGAATGACAAAGCTAGAGAGCTTGCAAAAAA	966

RESULT 7

AF224278

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

AF224278

1141 bp

mRNA

linear

PRI 18-JUL-2000

complete cds.

GI:9255808

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1141)

Xu, L.L., Shanmugam, N., Segawa, T., Seesterhenn, I. A., McLeod, D. G., Moul, J. W., and Srivastava, S.

A novel androgen-regulated gene, PMEPAL, located on chromosome 20q13 exhibits high level expression in prostate

Genomics 66 (3), 257-263 (2000)

2 (bases 1 to 1141)

Xu, L.L., Shanmugam, N., Segawa, T., Seesterhenn, I. A., McLeod, D. G., Moul, J. W., and Srivastava, S.

Direct Submission

Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street, Rockville, MD 20852, USA

Location/Qualifiers

1..1141

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="20"

/map="20q13.31-13.33"

/cell_line="LNCap"

1..1141

/gene="PMEPAL"

96..854

/gene="PMEPAL"

/note="type 1b transmembrane protein; expression is induced in response to the synthetic androgen, R1891; expression is abundant in, and restricted to, prostate"

glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP_004329"
/codon_start=1
/product="PMEPA1 protein"
/protein_id="AAP86322.1"
/db_xref="GI:9255809"
/translation="MAELFVQIIIIIVVMVWVVITCLSHYKLSARSFISHSQGRREDALSSBGLMPSTVSGNGIPVQVYAPRPTDLAVPPFAQRHFRFOPTYLYQHEIDLPTISLSDGEEPPYQGPCTQLRDPQQLNRESVRAPNRTIFDSDLMSARLGGPCPPSSNGSISATCYGSGRMGPPPTYSVIVHYPGSSFQHQSSGPPSLLEGTRLLHHTHIALPESAAIWSKEDKQKHPL"

ORIGIN

Query Match 87.3%; Score 926; DB 9; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 81 GTCTCTCGGAACACGAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG 140
Db 76 GTCTCTCGGAACACGAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG 135

Qy 141 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
Db 136 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195

Qy 201 CACGGTCTTTCATGACCGGACAGCCAGCGGCGGAGAGAGAGATGCTGCTGCTCAG 260
Db 196 CACGGTCTTTCATGACCGGACAGCCAGCGGCGGAGAGAGATGCTGCTGCTCAG 255

Qy 261 AAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 320
Db 256 AAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315

Qy 321 TCTAGCGCCCGCTCGGCGACACGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 380
Db 316 TCTAGCGCCCGCTCGGCGACACGACCGCTGCGCGCTGCGCGCTGCGCGCGGAGC 375

Qy 381 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACTGTCAGCAAGAGATGACCTGCGGCCA 440
Db 376 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACTGTCAGCAAGAGATGACCTGCGGCCA 435

Qy 441 CCATCTGCTGTCAGCGGAGAGGCCCCACCCCTACCGAGGCGCCCTGACCGCTCCAGC 500
Db 436 CCATCTGCTGTCAGCGGAGAGGCCCCACCCCTACCGAGGCGCCCTGACCGCTCCAGC 495

Qy 501 TTCGGGACCCCGAGCAGCAGCTGGAACCTGGAGTCCGTGCGCGACACCCCAACA 560
Db 496 TTCGGGACCCCGAGCAGCAGCTGGAACCTGGAGTCCGTGCGCGACACCCCAACA 555

Qy 561 GAACCATCTTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
Db 556 GAACCATCTTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615

Qy 621 GCAGTAATCTCGGCGATCAGCGCCACCTGTCAGCGAGCGGCGGCGATGAGGCGGCGC 680
Db 616 GCAGTAATCTCGGCGATCAGCGCCACCTGTCAGCGAGCGGCGGCGATGAGGCGGCGC 675

Qy 681 CGCCACCTTACAGCGAGTCTATCGGCCACTACCGGGGTCTCTTTCAGCAACAGCAGA 740
Db 676 CGCCACCTTACAGCGAGTCTATCGGCCACTACCGGGGTCTCTTTCAGCAACAGCAGA 735

Qy 741 GCAGTGGGCGCCCTCTCTTCTGAGAGGAGCGCGCTTCCACCAACACATCGCGGCC 800
Db 736 GCAGTGGGCGCCCTCTCTTCTGAGAGGAGCGCGCTTCCACCAACACATCGCGGCC 795

Qy 801 TAGAGCGCAGCCATCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGG 860
Db 796 TAGAGCGCAGCCATCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGG 855

Qy 861 GTCCCCAGGCGGCGCGGCTGCGGCTGCGTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAG 920
Db 856 GTCCCCAGGCGGCGCGGCTGCGGCTGCGTGTGTAAGGAGGAGGAGGAGGAGGAGGAGGAG 915

Qy 921 TAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 980
Db 916 TAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 975

Qy 981 CACCTCCCTGCTGATATAAATATTTACATGATGATGATGATGATGATGATGATGATGATG 1040
Db 976 CACCTCCCTGCTGATATAAATATTTACATGATGATGATGATGATGATGATGATGATGATG 1035

Qy 1041 GCTTGCACAAAAA 1057
Db 1036 GCTTGCACAAAAA 1052

RESULT 8
AR336830
LOCUS AR336830 1140 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 1 from patent US 6566130.
ACCESSION AR336830
VERSION AR336830.1 GI:33722680
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Srivastava,S.; Moul,J.W.; Xu,L.L. and Segawa,T.
TITLE Androgen-regulated gene expressed in prostate tissue
JOURNAL Patent: US 6566130-A 1 20-MAY-2003;
FEATUES Location/Qualifiers
source 1..1140
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 86.0%; Score 912; DB 6; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 95 CAGGCAATGGCGGAGCTGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 154
Db 89 CAGGCAATGGCGGAGCTGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 148

Qy 155 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 214
Db 149 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 208

Qy 215 AGCGCGCACAGCCAGGCGGAGGAGAGATGCTGCTCTCAGAGAGATGCTGCTGCTGCTGCTG 274
Db 209 AGCGCGCACAGCCAGGCGGAGGAGAGATGCTGCTCTCAGAGAGATGCTGCTGCTGCTGCTG 268

Qy 275 CCCTCGGAGAGACAGTGTTCAGGCAACCGAATCCAGAGCCGCGAGCTCTACGCGCGCCT 334
Db 269 CCCTCGGAGAGACAGTGTTCAGGCAACCGAATCCAGAGCCGCGAGCTCTACGCGCGCCT 328

Qy 335 CGGCGCACCGGCGCTGCGCGCTGCGCGCTTTCGCGCGAGCGAGCGCTTTCACACCGCTTC 394
Db 329 CGGCGCACCGGCGCTGCGCGCTGCGCGCTTTCGCGCGAGCGAGCGCTTTCACACCGCTTC 388

Qy 395 CAGCCCACTTACCTGCTACCTGACGACGAGATGACCTGCGCGCCCAACATCTGCTGTCA 454
Db 389 CAGCCCACTTACCTGCTACCTGACGACGAGATGACCTGCGCGCCCAACATCTGCTGTCA 448

Qy 455 GACGGGAGGAGCGCCCAACCTTACGAGGCGCGCTGCACTTCCAGCTTGGGAGCCCGAG 514
Db 449 GACGGGAGGAGCGCCCAACCTTACGAGGCGCGCTGCACTTCCAGCTTGGGAGCCCGAG 508

Qy 515 CAGCAGCTGGAACCTGAAACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 574
Db 509 CAGCAGCTGGAACCTGAAACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 568

Qy 575 AGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 634
Db 569 AGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 628

[illegible]

47		502	CTGATGTTAGTGCAGGCTGGGGGGCCCCCTGCCCCACAGCACTAACTCGSSCATCAGC	545
Db		486	CTGATGTTAGTGCAGGCTGGGGGGCCCCCTGCCCCACAGCACTAACTCGSSCATCAGC	545
Qy		641	GCCACGTGCTACGGCAGCGCGGGCGCATATGGAGGGGGCGCGCCACCTACAGCAGATC	700
Db		546	GCCACGTGCTACGGCAGCGCGGGCGCATATGGAGGGGGCGCGCCACCTACAGCAGATC	605
Qy		701	ATCGGCCACTACCGGGGTCTCTTCTTCAGCACACGACAGCAGTGGGCGCGCCCTCTCTTG	760

Db	606	ATCGGCGCACTACCGGGTCTCTCTCCAGCACACAGAGAGCTGTGGCGCGCCCTCTCTTG	666
Qy	761	CTGAGGGGACCGGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGGACGCCATCTGG	820
Db	666	CTGAGGGGACCGGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGAGCCCATCTGG	725
Qy	821	AGCAAGAGAGGATAAACAGAAAGGACACCTCTCTTAGGGTCCCGAGGGGGCGCGGCT	880
Db	726	AGCAAGAGAGGATAAACAGAAAGGACACCTCTCTTAGGGTCCCGAGGGGGCGCGGCT	785
Qy	881	GGGCTCGGTAGGTGAAGAGCGCAACACTCGCGCTTCTTAGAAGAGGAGGTGAGAGGAA	940
Db	786	GGGCTCGGTAGGTGAAGAGCGCAACACTCGCGCTTCTTAGAAGAGGAGGTGAGAGGAA	845
Qy	941	GGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCGCTCCCACTCCCTGTGTATAATA	1000
Db	846	GGCGGGGGGGCAGCAACGATCGTGTGGCCCTCCGCTCCCACTCCCTGTGTATAATA	905

Db 906 TTTACATGTGATGTCGGTCTGATGCACACAGCTTAGAGAGCTTGCAAAAAAAAAA 965

Qy 1061 A 1061
|
966 A 966

Db 966 A 966

RESULT 10			
BD272516			
LOCUS	BD272516	969 bp	DNA linear PAT 17-JUL-2000

Qy		641	GCCACGTGCTACCGCAGCGGGCGGCATGAGGGGGCGCGGCCACCTTACAGCGAGGTC	700
Dd		546	GCCACGTGCTACCGCAGCGGGCGGCATGAGGGGGCGCGGCCACCTTACAGCGAGGTC	605
Qy		701	ATCGGGCACTACCGGGGCTCTCTTTCCAGCACACAGCAGCAGTAGTGCGCGCCCTCTCTTG	760
Dd		606	ATCGGGCACTACCGGGGCTCTCTTTCCAGCACACAGCAGCAGTAGTGCGCGCCCTCTCTTG	665
Qy		761	CTGAGGGGACC CGGCTTCCACACACACATCGCGCCCTTAGAGAGCGGAGCATCTGG	820
Dd		666	CTGAGGGGACC CGGCTTCCACACACATCGCGCCCTTAGAGAGCGGAGCATCTGG	725
Qy		821	AGCAAAGAGAAGCATAAACAGAAAGCACAACCTCTCTTAGGGTCCCAGGGGGGCGCGGCT	880
Dd		726	AGCAAAGAGAAGCATAAACAGAAAGCACAACCTCTCTTAGGGTCCCAGGGGGGCGCGGCT	785
Qy		881	GGGGCTGCTAGGTGAAGGAGGAGAACATCTCGCGCTCTTTAAGAGAGGAGTAGAGAGAA	940
Dd		786	GGGGCTGCTAGGTGAAGGAGGAGAACATCTCGCGCTCTTTAAGAGAGGAGTAGAGAGAA	845
Qy		941	GGGGGGGGCGCAGCAACCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATA	1000
Dd		846	GGGGGGGGCGCAGCAACCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATA	905
Qy		1001	TTTTACATGTGATGTCTGTGCTGTAATGCACAGCTAAGAGAGCTTGCAAAAA	1060
Dd		906	TTTTACATGTGATGTCTGTGCTGTAATGCACAGCTAAGAGAGCTTGCAAAAA	965
Qy		1061	A 1061	
Dd		966	A 966	

RESULT 11
AY128643

LOCUS Homo sapiens PMP21 variant A protein mRNA linear PRI 10-APR-2003
DEFINITION AY128643
ACCESSION AY128643
VERSION AY128643.1 GI:22121998
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E., Willson,J.K.V., Lu,S., Nosrati,A., Swinler,S., Beard,L., Lutterbaugh,J.D., Willis,J., Platzner,P. and Markowitz,S.
TITLE PMP21, a transforming growth factor-beta-induced marker of terminal colonocyte differentiation whose expression is maintained in primary and metastatic colon cancer
JOURNAL Cancer Res. 63 (7), 1568-1575 (2003)
MEDLINE 22557253
PUBMED 12670906

REFERENCE
AUTHORS Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E., Willson,J.K.V., Lu,S., Nosrati,A., Swinler,S., Beard,L., Lutterbaugh,J.D., Willis,J., Platzner,P. and Markowitz,S.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2002) Department of Medicine, Case Western Reserve University/Howard Hughes Medical Institute, 11001 Cedar Ave., Cleveland, OH 44106, USA

FEATURES
source location/Qualifiers
1..1818 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.31-13.33"
145..858 /codon_start=1
product="PMP21 variant A protein"
protein_id="AA09277.1"

CDS

/db_xref="GI:22121999"		/db_xref="GI:22121999"			
/translation="MMVMVVITCLSHYKLSARFISHSQRRRDLSEGLMPP SESTVSGNGIPFPQVTPDRDLAVPPAQRFRHFRPTPYLQHEIDLPPTISL SDGEPFPYQCTQLRDPQOELNRESVRAPNRTIFDLDLMDSLRGCPSPSS NSGIGATCYGSGRMGGPPPTTSEVIGHYFGSSFGHQSSGPPSLLEGLRHLHTIAP LESAAALSKKDKQKHPL"		/translation="MMVMVVITCLSHYKLSARFISHSQRRRDLSEGLMPP SESTVSGNGIPFPQVTPDRDLAVPPAQRFRHFRPTPYLQHEIDLPPTISLSDGEPFP GIPEQVTPDRDLAVPPAQRFRHFRPTPYLQHEIDLPPTISLSDGEPFP YQPCCTQLRDPQOELNRESVRAPNRTIFDLDLMDSLRGCPSPSPSSNSGIGATC YSGRMGGPPPTTSEVIGHYFGSSFGHQSSGPPSLLEGLRHLHTIAPLESAAALIS			
ORIGIN		ORIGIN			
Query Match 85.1%; Score 903; DB 9; Length 1818; Best Local Similarity 99.9%; Pred. No. 0; Matches 953; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		Query Match 85.0%; Score 902; DB 9; Length 4839; Best Local Similarity 99.9%; Pred. No. 0; Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	104	GGCGAGCTGAGCTTCTTCAGATCATCATCGTGGTGATGATGCTGATGCTG 163	QY	105	CGAGCTGAGCTTCTTCAGATCATCATCGTGGTGATGATGCTGATGCTG 164
DB	103	GGCGAGCTGAGCTTCTTCAGATCATCATCGTGGTGATGATGCTGATGCTG 162	DB	430	CGAGCTGAGCTTCTTCAGATCATCATCGTGGTGATGATGCTGATGCTG 489
QY	164	GTGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTGACGGTCTTCATCAGCCGAC 223	QY	165	TGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTGACGGTCTTCATCAGCCGAC 224
DB	163	GTGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTGACGGTCTTCATCAGCCGAC 222	DB	490	TGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTGACGGTCTTCATCAGCCGAC 549
QY	224	AGCAGGGGCGGAGGAGAGATGCTTCTCTCAGAGGATGCTGTGGCTCTGGAG 283	QY	225	GCAGGGGCGGAGGAGAGATGCTTCTCTCAGAGGATGCTGTGGCTCTGGAG 284
DB	223	AGCAGGGGCGGAGGAGAGATGCTTCTCTCAGAGGATGCTGTGGCTCTGGAG 282	DB	550	GCAGGGGCGGAGGAGAGATGCTTCTCTCAGAGGATGCTGTGGCTCTGGAG 609
QY	284	AGCACAGTGTTCAGCAACGGAATCCAGAGCGCAGGCTCTACGCCGCTCGGCCACC 343	QY	285	GCACAGTGTTCAGCAACGGAATCCAGAGCGCAGGCTCTACGCCGCTCGGCCACC 344
DB	283	AGCACAGTGTTCAGCAACGGAATCCAGAGCGCAGGCTCTACGCCGCTCGGCCACC 342	DB	610	GCACAGTGTTCAGCAACGGAATCCAGAGCGCAGGCTCTACGCCGCTCGGCCACC 669
QY	344	GACCGCTGCGGCTGCGGCTTCTGCGCCAGCGGGAGGCTTCCACCGCTTCCAGCCACC 403	QY	345	ACCGCTGCGGCTGCGGCTTCTGCGCCAGCGGGAGGCTTCCACCGCTTCCAGCCACC 404
DB	343	GACCGCTGCGGCTGCGGCTTCTGCGCCAGCGGGAGGCTTCCACCGCTTCCAGCCACC 402			
QY	404	TATCCGTACCTGACGACGAGATCGACTGCGCGCCCAACATCTCGCTGTGACAGGGGAG 463			
DB	403	TATCCGTACCTGACGACGAGATCGACTGCGCGCCCAACATCTCGCTGTGACAGGGGAG 462			
QY	464	GAGCCCCACCTTACAGAGGCGCTGACCTCCAGCTTGGGACCCCGAGCAGCAGCTG 523			
DB	463	GAGCCCCACCTTACAGAGGCGCTGACCTCCAGCTTGGGACCCCGAGCAGCAGCTG 522			
QY	524	GAACTGAACCGGAGTCTGCTGCGCGCACCCGCAACCAACATCTTCCAGCAGTACCTG 583			
DB	523	GAACTGAACCGGAGTCTGCTGCGCGCACCCGCAACCAACATCTTCCAGCAGTACCTG 582			
QY	584	ATGATAGTGTTCAGGCTGGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 643			
DB	583	ATGATAGTGTTCAGGCTGGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 642			
QY	644	ACGTGCTACGGCAGCGCGCGCGCATGAGGGGCGCGCGCGCGCGCGCGCGCGCGCG 703			
DB	643	ACGTGCTACGGCAGCGCGCGCGCATGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 702			
QY	704	GCCACTTACCGGGTCTCTCTTCCAGCACGAGCAGTGGCGCGCGCTCTCTGCTG 763			
DB	703	GCCACTTACCGGGTCTCTCTTCCAGCACGAGCAGTGGCGCGCGCTCTCTGCTG 762			
QY	764	GAGGGACCGGCTTCCACACACATCGCGCGCTTACAGAGGAGGAGCATCTGGAGC 823			
DB	763	GAGGGACCGGCTTCCACACACATCGCGCGCTTACAGAGGAGGAGCATCTGGAGC 822			
QY	824	AAGAGGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCTGG 883			
DB	823	AAGAGGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCTGG 882			
QY	884	GCTGCTAGGTGAAAGGACAGAACTCTCCGCGCTTCTTAGAGAGGAGTGAAGGAGGC 943			
DB	883	GCTGCTAGGTGAAAGGACAGAACTCTCCGCGCTTCTTAGAGAGGAGTGAAGGAGGC 942			
QY	944	GGGGGCGCAGCAACGATCGTGTGGCTTCCCTCTCCACCTCTGCTGTATTAATTTT 1003			
DB	943	GGGGGCGCAGCAACGATCGTGTGGCTTCCCTCTCCACCTCTGCTGTATTAATTTT 1002			
QY	1004	ACATGTGATGCTGCTGATGCAAGCTAAGAGAGCTTGGCAAAAAA 1057			

Db	1003	ACATGTGATGCTGGTCTCAATGCACACGCTAAGAGCTTGGCAAAAAA	1056
RESULT 12			
LOCUS	AF305616	4839 bp	mRNA linear PRI 21-OCT-2001
DEFINITION	Homo sapiens STAG1/PMEPAL mRNA, complete cds.		
ACCESSION	AF305616		
VERSION	AF305616.1 GI:16303741		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4839)		
TITLE	Rae,P.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.		
JOURNAL	Characterization of a novel gene, STAG1/PMEPAL, upregulated in renal cell carcinoma and other solid tumors		
MEDLINE	Mol. Carcinog. 32 (1), 44-53 (2001)		
PUBMED	21453682		
REFERENCE	2 (bases 1 to 4839)		
AUTHORS	Rae,P.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-SEP-2000) Centre for Molecular Biotechnology, Queensland University of Technology, 2 George St, Brisbane, QLD 4001, Australia		
FEATURES	Location/Qualifiers		
source	1..4839		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="20"		
	/map="20q13.2-q13.33"		
	321..1184		
	/note="unknown function"		
	/codon_start=1		
	/product="STAG1/PMEPAL"		
	/protein_id="AA116781.1"		
	/db_xref="GI:16303742"		
	/translation="MRLMGVNSTAAAAAGQPNVCTCNKRSIFQSMBITELEFVQI IIIVVMVVVITCLSHYKLSARFISHSQRRRDLSEGLMPPSESTVSGN GPEPOVAPRDLAVPPAQRFRHFRPTPYLQHEIDLPPTISLSDBEP YGPCTQLRDPQOELNRESVRAPNRTIFDLDLMDSLRGCPSPSSNGISATC YSGGMRMBGPPPTTSEVIGHYFGSFGHQSSGPPSLLEGLRHLHTIAPLESAAINS KERDKQKHPL"		
	polyA_signal 2158..2163		
	polyA_signal 2463..2468		
	polyA_signal 4818..4823		
ORIGIN			
Query Match	85.0%; Score 902; DB 9; Length 4839;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 952; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	105	CGGAGCTGGAGTTGTTGTCAGATCATCATCTGCTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	

[illegible]

Qy	345	ACGCGCTGGCGGTGCGCGCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCT	404
Db	762		821
Qy	405	ATCGGTACCTGCGAGCAGAGATCGACCTGCGCGCCACCATCTCGCTGTCTCAGACCGGGAGG	464
Db	822		881
Qy	465	AGCGCCACCCCTTACCGAGGCGCTTGCACCTTCCAGCTTCCGAGACCGCGAGCAGCTGG	524
Db	882		941
Qy	525	AACCTGACCGGGAGTCCGGTGGCGCACCCGCCAAGAACCATCTTCGACAGTGAACCTGA	584
Db	942		1001
Qy	585	TGGATAGTGCAGGCTGGGGGGCGCTTCCGCCAGCAGTAACCTCGGGCATCAGCGCCA	644
Db	1002		1061
Qy	645	CTGTCTACGGCAGCGGGGGCGATGAGAGGGGGCGCGCGCCACCTACAGCGAGGTATCG	704
Db	1062		1121
Qy	705	GCCACTACCGGGGTCTCTTCCAGCACCGAGCAGTGGGGCGCGCGCTTCTTGGCTGG	764
Db	1122		1181
Qy	765	AGGGGACCCCGCTCCACACACACATCGCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA	824
Db	1182		1241
Qy	825	AAGAGAAGGATAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG	884
Db	1242		1301
Qy	885	CTCGGTAGGTGAAGGCAG	904
Db	1302		1321

Search completed: May 26, 2004, 19:11:33
Job time : 6859 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: May 26, 2004, 14:52:22 ; Search time 667 Seconds
(without alignments)
6757.628 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tctccttggtggtggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO NUC
Gapop_60.0, Gapext 60.0

Searched: 3373863 seqs, 212499041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1061	100.0	1061	3 AAA47429	AAA47429 Sequence
2	1057	99.6	1085	9 ADC37324	ADC37324 Nuclear f
3	1056	99.5	1334	7 ABZ36103	ABZ36103 Human sec
4	961	90.6	969	3 AAA75151	AAA75151 cDNA enco
5	953	89.8	1383	9 ADC37326	ADC37326 Nuclear f
6	952	89.7	969	3 AAA75163	AAA75163 cDNA clon
7	926	87.3	1140	6 ABK92120	ABK92120 Prostate
8	926	87.3	1141	9 ADB75588	ADB75588 Prostate
9	926	87.3	1850	7 ACC49536	ACC49536 Tumour-as
10	912	86.0	1140	9 AAD60105	AAD60105 Human and
11	910	85.8	969	3 AAA75164	AAA75164 cDNA clon
12	910	85.8	969	3 AAA75165	AAA75165 cDNA clon
13	902	85.0	4839	7 ACC49552	ACC49552 Tumour-as
14	891	84.0	1066	4 AAI57868	AAI57868 Human pol
15	788	74.3	806	7 ACC49537	ACC49537 Tumour-as
16	749	70.6	1321	6 ABK12137	ABK12137 Human cdn
17	548	51.6	1069	4 AAI59654	AAI59654 Human pol
18	509	48.0	1583	6 ABS61424	ABS61424 Prostate
19	415	39.1	474	7 ABZ84732	ABZ84732 Toxicolog
20	350	33.0	408	5 AAF65983	AAF65983 Novel hum
21	322	30.3	812	2 AAZ52964	AAZ52964 Human pro
22	313	29.5	693	6 ABK12143	ABK12143 Human MIV
23	269	25.4	467	8 ACH14862	ACH14862 Human adu

ALIGNMENTS

RESULT 1

AAA47429 standard; DNA; 1061 BP.

AC AAA47429;

DT 20-OCT-2000 (first entry)

DE Sequence encoding human neuron-associated protein.

XX Neuron associated protein; NEUAP; neurological disorder; epilepsy;
ischemic cerebrovascular disease; stroke; cerebral neoplasm;
Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
Parkinson's disease; demyelinating disease; meningitis; prion disease;
Kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;
muscular dystrophy; central nervous system; CNS;
peripheral nervous system; PNS; myopathy; schizophrenia;
actinic keratosis; arteriosclerosis; atherosclerosis; burkitt;
cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
autoimmune disease; inflammation; acquired immunodeficiency syndrome;
AIDS; Addison's disease; adult respiratory distress syndrome; allergy;
ankylosing spondylitis; amyloidosis; anaemia; asthma;
Werner syndrome; trauma; human; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

CDS 101..859
/*tag= a
/product= "Neuron associated protein"

XX WO200034477-A2.

PD 15-JUN-2000.

XX 10-DEC-1999; 99WO-US030408.

PR 11-DEC-1998; 98US-00210083.

PR 09-FEB-1999; 99US-0119365P.

PR 16-MAR-1999; 99US-0124687P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;

PI Yang J, Lu DAM, Azimzai Y;

XX PA (ASAH) ASAH KASEI KK.
 XX PI Matsuda A, Muramatsu S;
 XX DR WPI; 2003-505282/47.
 XX DR P-PSDB; ADC37325.
 XX PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 XX PT useful for treating inflammation, autoimmune diseases, cancers,
 XX PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 XX PT ischemic disorders.
 XX PS Claim 4; SEQ ID NO 157; 938pp; English.
 XX CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX SQ Sequence 1085 BP; 223 A; 352 C; 334 G; 176 T; 0 U; 0 Other;
 Query Match 99.6%; Score 1057; DB 9; Length 1085;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCCTCTCTGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCGAGTGC 60
 DB 28 TCCTCTCTGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCGAGTGC 87
 QY 61 TGGGAAGTCTGAAGCGGACGCTCTCTCGGAAACAGGCAATGGCGAGCTGGAGTTGT 120
 DB 88 TGGGAAGTCTGAAGCGGACGCTCTCTCGGAAACAGGCAATGGCGAGCTGGAGTTGT 147
 QY 121 TCAGATCATCATCATCTGGTGGTGGTGAATGATGGTGGTGGTGGTGGTGGTGGT 180
 DB 148 TCAGATCATCATCATCTGGTGGTGGTGAATGATGGTGGTGGTGGTGGTGGTGGT 207
 QY 181 GAGCCACTAGACTGTCTGACCGTCTTCATCAGCGGCACACGCGGCGGCGGAG 240
 DB 208 GAGCCACTAGACTGTCTGACCGTCTTCATCAGCGGCACACGCGGCGGCGGAG 267
 QY 241 AGAAGATGCCCTGCTCTCAGAAAGATGCTGTGGCCCTCGAGAGCAGTGTTCAGGCA 300
 DB 268 AGAAGATGCCCTGCTCTCAGAAAGATGCTGTGGCCCTCGAGAGCAGTGTTCAGGCA 327
 QY 301 CGGAATCCGAGCGCGAGGTCTACGCGCCCGCTCGGCCACCGACCGCTGGCGGTGCC 360
 DB 328 CGGAATCCGAGCGCGAGGTCTACGCGCCCGCTCGGCCACCGACCGCTGGCGGTGCC 387
 QY 361 GCCCTTCGCGCGAGCGGTCTCCACCGCTTCAGCGCCACCTATCGTACCTGCAGCA 420
 DB 388 GCCCTTCGCGCGAGCGGTCTCCACCGCTTCAGCGCCACCTATCGTACCTGCAGCA 447
 QY 421 CGAGATCGACCTGCGCGCCACATCTCGCTGTGAGCGGAGCGAGCGCCACCTACCA 480
 DB 448 CGAGATCGACCTGCGCGCCACATCTCGCTGTGAGCGGAGCGAGCGCCACCTACCA 507
 QY 481 GGGCGCCCTGCAACCTTCAGCTGTGGGACCCCGAGGAGCAGTGTGAACCTGAACCGGAGTC 540
 DB 508 GGGCGCCCTGCAACCTTCAGCTGTGGGACCCCGAGGAGCAGTGTGAACCTGAACCGGAGTC 567
 QY 541 GGTGGCGGCGCCCGAAGCAACCTTCGACAGTGAACCTGATGATGATGATGATGATGAT 600
 DB 568 GGTGGCGGCGCCCGAAGCAACCTTCGACAGTGAACCTGATGATGATGATGATGATGAT 627
 QY 601 GGGCGGCGCCCTGCGCGCCCGGAGTGAACCTCGGGGATCAGCGCCACCTGCTACGCGAGCGG 660
 DB 628 GGGCGGCGCCCTGCGCGCCCGGAGTGAACCTCGGGGATCAGCGCCACCTGCTACGCGAGCGG 687
 QY 661 CGGCGGCGATGAGGGGCGCGCGCCCGACCTACAGCGAGGTGATCGGCGCACTACCGGGGTC 720

DB 688 CGGCGCATGAGGGGCGCGCCCTACAGCGAGGTCACTCGGCCACTACCGGGGTC 747
 QY 721 CTCCTTCCAGCACCGAGGAGCGAGTGGGGCGCGCCCTCTCTGCTGAGGGGAGCCCGGCTCCA 780
 DB 748 CTCCTTCCAGCACCGAGGAGCGAGTGGGGCGCGCCCTCTCTGCTGAGGGGAGCCCGGCTCCA 807
 QY 781 CCACACACATCGCGCCCTTAGAGAGCGCGAGCCATCTGGAGCAAGAGAGGATAAACA 840
 DB 808 CCACACACATCGCGCCCTTAGAGAGCGCGAGCCATCTGGAGCAAGAGAGGATAAACA 867
 QY 841 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCTGGGGCTGCGTAGTGTGAAAG 900
 DB 868 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGCGGCTGGGGCTGCGTAGTGTGAAAG 927
 QY 901 CGAGAACACTCGCGCTCTTGAAGAGCGAGTGAAGAGGAGCGGGGGCGCGAGCAACGC 960
 DB 928 CGAGAACACTCGCGCTCTTGAAGAGCGAGTGAAGAGGAGCGGGGGCGCGAGCAACGC 987
 QY 961 ATCGTGTGGCCCTCCCGTCCACCTCCCTGTTGTTATATAATATTTACATGTGATGTCTGGTC 1020
 DB 988 ATCGTGTGGCCCTCCCGTCCACCTCCCTGTTGTTATATAATATTTACATGTGATGTCTGGTC 1047
 QY 1021 TGAATGCAAGCTTAAGAGAGCTTGCACCAAAAAA 1057
 DB 1048 TGAATGCAAGCTTAAGAGAGCTTGCACCAAAAAA 1084
 RESULT 3
 ABZ36103
 ID ABZ36103 standard; cDNA; 1334 BP.
 XX AC ABZ36103;
 XX AC
 DT 10-FEB-2003 (first entry)
 XX DT
 DE Human secretory polynucleotide SPTM SEQ ID NO 267.
 XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianemic; anti-HIV; human immunodeficiency virus;
 XX secretory polynucleotide; secretory protein; gene; ss.
 OS Homo sapiens.
 XX OS
 PN WO200283876-A2.
 XX PN
 PD 24-OCT-2002.
 XX PD
 PF 27-MAR-2002; 2002WO-US009921.
 XX PF
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX (INCYTE) INCYTE GENOMICS INC.
 XX PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshley SR;
 XX PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX WIPI; 2003-075543/07.
 DR

XX	PA	(MILL-) MILLENNIUM PHARM INC.	Db	426	CTGGAACGACGGGAGTGGTGGCGGCGACCCGCCAAACAGAACCATCTTCGACAGTGAC	485
XX	PI	Barnes TM, Holtzman DA, Sharp JD, Fraser CC;	Qy	581	CTGATGATAGTGCAGGCTGGGGCGGCGCCCTGCGCCCGCCAGCAGTAACCTCGGGCATCAGC	640
XX	XX	P-PSDB; AAB18449.	Db	486	CTGATGATAGTGCAGGCTGGGGCGGCGCCCTGCGCCCGCCAGCAGTAACCTCGGGCATCAGC	545
XX	DR	WPI; 2000-579269/54.	Qy	641	GCCACGTGCTACCGCGGCGGCGGCGCATGAGGGGGCGCGCCACCTACAGCGAGGTC	700
XX	XX	Novel human and murine secreted proteins designated TANGO 216, 261, 262,	Db	546	GCCACGTGCTACCGCGGCGGCGGCGCATGAGGGGGCGCGCCACCTACAGCGAGGTC	605
XX	PT	266 and 267 useful as modulating agents of cellular processes, e.g. for	Qy	701	ATCGGCCACTACCGCGGCGGCGGCGCATGAGGGGGCGCGCCACCTACAGCGAGGTC	760
XX	PT	treating cancer.	Db	606	ATCGGCCACTACCGCGGCGGCGGCGCATGAGGGGGCGCGCCACCTACAGCGAGGTC	665
XX	PS	Claim 2; Fig 5; 175pp; English.	Qy	761	CTGAGGGGACCGCGGCTCCACCAACACACATCGCGCCCTAGAGAGCGGAGCCATCTGG	820
XX	CC	The present sequence encodes a human TANGO 261 polypeptide. The	Db	666	CTGAGGGGACCGCGGCTCCACCAACACACATCGCGCCCTAGAGAGCGGAGCCATCTGG	725
XX	CC	specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO	Qy	821	AGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT	880
XX	CC	267. The TANGO polypeptides can be used to modulate cellular	Db	726	AGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT	785
XX	CC	adhesion. The proteins can be used to treat any von Willebrand factor-	Qy	881	GCGGCTCGGTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	940
XX	CC	associated disorder, regulate extracellular matrix structuring, cellular	Db	786	GCGGCTCGGTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	845
XX	CC	adhesion, and cell trafficking and/or migration, modulate cellular	Qy	941	GCGGCGGGCGGCGAGCAAGCATCGTGTGGCGCCCTCCACCTCCCTGCTGTATTAATA	1000
XX	CC	interactions, modulate cell adhesion in proliferative disorders, such as	Db	846	GCGGCGGGCGGCGAGCAAGCATCGTGTGGCGCCCTCCACCTCCCTGCTGTATTAATA	905
XX	CC	cancer, modulate the proliferation, differentiation, and/or function of	Qy	1001	TTTACATGTGATGCTGGTCTGAATGCAAGCTTAAGAGGCTTGCAAAAAA	1060
XX	CC	cells that appear in the bone marrow, and leukocytes, treat bone marrow,	Db	906	TTTACATGTGATGCTGGTCTGAATGCAAGCTTAAGAGGCTTGCAAAAAA	965
XX	CC	blood and hematopoietic associated diseases and disorders, atelectasis,	Qy	1061	A 1061	
XX	CC	asthma and bronchiectasis, intestinal disorders, spleen associated	Db	966	A 966	
XX	CC	diseases, modulate renal disorders, treat cardiovascular disorders such				
XX	CC	as ischemic heart disease, modulate the proliferation, differentiation,				
XX	CC	and/or function of bone and cartilage cells and to treat bone and/or				
XX	CC	cartilage associated diseases or disorder. They may also be used to treat				
XX	CC	disorders associated with the ovaries, cerebral oedema, hydrocephalus,				
XX	CC	brain herniations, iatrogenic disease, inflammations, bacterial and viral				
XX	CC	meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's				
XX	CC	disease, multiple sclerosis, brain cancers, hydrocephalus and				
XX	CC	encephalitis, and treat hepatic disorders				
XX	SQ	Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 U; 0 Other;				
Query Match 90.6%; Score 961; DB 3; Length 969;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	101	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCTGTTGATGATGATGATGATG	160			
Db	6	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCTGTTGATGATGATGATGATG	65			
Qy	161	GTGGTGATCATGTCCTGCTGAGCCACTACAGCTGTCTGCAAGTCTTTCATCAGCCGG	220			
Db	66	GTGGTGATCATGTCCTGCTGAGCCACTACAGCTGTCTGCAAGTCTTTCATCAGCCGG	125			
Qy	221	CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGATGCTGTGGCCCTCG	280			
Db	126	CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGATGCTGTGGCCCTCG	185			
Qy	281	GAGAGCACAGTGTGAGGCAACGGAATCCAGAGCCGCGAGGTCTACCGCCCGCTCGGCC	340			
Db	186	GAGAGCACAGTGTGAGGCAACGGAATCCAGAGCCGCGAGGTCTACCGCCCGCTCGGCC	245			
Qy	341	ACCGACCGCTGGCGCTGGCGCCCTTGGCGCCAGCGGGAGCGCTTCACCGCTTCAGGCC	400			
Db	246	ACCGACCGCTGGCGCTGGCGCCCTTGGCGCCAGCGGGAGCGCTTCACCGCTTCAGGCC	305			
Qy	401	ACCTATCCCTACTCTGACAGCACAGATGCGACCTGCGCCGCCACCATCTCGGTGTGACAGCG	460			
Db	306	ACCTATCCCTACTCTGACAGCACAGATGCGACCTGCGCCGCCACCATCTCGGTGTGACAGCG	365			
Qy	461	GAGGAGCCCGCCACCTTACAGGCGCCCTGACCTTCAGCTTCGGGACCCCGAGCAGCAG	520			
Db	366	GAGGAGCCCGCCACCTTACAGGCGCCCTGACCTTCAGCTTCGGGACCCCGAGCAGCAG	425			
Qy	521	CTGGAACGACGGGAGTGGTGGCGGCGACCCGCCAAACAGAACCATCTTCGACAGTGAC	580			

RESULT 5

ADC37326
ID ADC37326 standard; DNA; 1383 BP.

XX AC ADC37326;

XX AC AC

DT 18-DEC-2003 (first entry)

XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 159.

XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
cancer; infectious disease; bone disease; AIDS;

XX neurodegenerative disease; ischaemic disorder; Antiinflammatory;

XX Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

XX Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

XX Homo sapiens.

OS WO2003048202-A2.

PN 12-JUN-2003.

XX 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

XX (ASAH) ASAH KASBI KK.

XX Matsuda A, Muramatsu S;

PI WPI; 2003-505282/47.

XX

XX

XX

DB
976 CACCTCCCGTGTGTAATAATATTACATGTGGTAGTGTCAGGCCTGMAAGCACHACAGCTAAGA 103

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 201 CACGGTCTTTCATCAGCGCGCACAGCGAGGCGCGGAGAGAGATGCCCTGTCTCTCAG 260
 Db 196 CACGGTCTTTCATCAGCGCGCACAGCGAGGCGCGGAGAGAGATGCCCTGTCTCTCAG 255
 QY 261 AAGATGCTGTGGCTCTCGAGAGGCAAGTGTAGGGAACGGAATCCAGAGCGCAGG 320
 Db 256 AAGATGCTGTGGCTCTCGAGAGGCAAGTGTAGGGAACGGAATCCAGAGCGCAGG 315
 QY 321 TCTAGCGCCCGCTCGCGCCACAGCGCGCTGGCGCTGGCGCTTCCGCCAGCGCGAGC 380
 Db 316 TCTAGCGCCCGCTCGCGCCACAGCGCGCTGGCGCTGGCGCTTCCGCCAGCGCGAGC 375
 QY 381 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 440
 Db 376 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 435
 QY 441 CCACTCTGCTGTGAGAGCGGAGAGGCGCCACCTACCGAGGCGCTTCCGCCAGCTCAGC 500
 Db 436 CCACTCTGCTGTGAGAGCGGAGAGGCGCCACCTACCGAGGCGCTTCCGCCAGCTCAGC 495
 QY 501 TCCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGCGGCGACCCCAACA 560
 Db 496 TCCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGCGGCGACCCCAACA 555
 QY 561 GAACCATCTTCCAGCTGACCTGATGATAGTGCAGGCTGGGCGGCGCTTCCGCCAGC 620
 Db 556 GAACCATCTTCCAGCTGACCTGATGATAGTGCAGGCTGGGCGGCGCTTCCGCCAGC 615
 QY 621 GCAGTAACCTCGGCGATCAGCGCCACCTGCTACGCGAGCGGCGCGCATGAGGCGCGC 680
 Db 616 GCAGTAACCTCGGCGATCAGCGCCACCTGCTACGCGAGCGGCGCGCATGAGGCGCGC 675
 QY 681 CGCCCACTACAGGAGTCACTGGCCCACTACCGGGGTCTCTTCCAGCACAGCAGA 740
 Db 676 CGCCCACTACAGGAGTCACTGGCCCACTACCGGGGTCTCTTCCAGCACAGCAGA 735
 QY 741 GCAGTGGCGCGCTCTCTGCTGAGGAGGACCGCGCTCCACACACACATCGCGCCCC 800
 Db 736 GCAGTGGCGCGCTCTCTGCTGAGGAGGACCGCGCTCCACACACACATCGCGCCCC 795
 QY 801 TAGAGAGCGCAGCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGG 860
 Db 796 TAGAGAGCGCAGCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGG 855
 QY 861 GTCCCGAGGGGCGCGGCTGGGCTGCTGAGTGAAGAGGAGCACTCCGCGCTTCT 920
 Db 856 GTCCCGAGGGGCGCGGCTGGGCTGCTGAGTGAAGAGGAGCACTCCGCGCTTCT 915
 QY 921 TAGAGAGGAGTGAAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 980
 Db 916 TAGAGAGGAGTGAAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 975
 QY 981 CACTCTCTGTATATAATTTACATGTGATGTCTGGTCTGAATGCAAGCTTAAGAGA 1040
 Db 976 CACTCTCTGTATATAATTTACATGTGATGTCTGGTCTGAATGCAAGCTTAAGAGA 1035
 QY 1041 GCTTGCAGAAAAA 1057
 Db 1036 GCTTGCAGAAAAA 1052

RESULT 9
 ACC49536
 ID ACC49536 standard; cDNA; 1850 BP.

XX AC ACC49536;

XX AC

DT 01-JUL-2003 (first entry)

DE Tumour-associated antigenic target protein TAT180 cDNA SEQ ID NO:44.
 XX Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
 KW cancer; gene; ss.

XX Homo sapiens.
 OS WO2003024392-A2.
 XX 27-MAR-2003.
 XX 11-SEP-2002; 2002WO-US028859.
 XX 18-SEP-2001; 2001US-0323268P.
 PR 19-OCT-2001; 2001US-0339227P.
 PR 07-NOV-2001; 2001US-0336827P.
 PR 20-NOV-2001; 2001US-0331906P.
 PR 02-JAN-2002; 2002US-0345444P.
 PR 03-APR-2002; 2002US-0369724P.
 PR 19-AUG-2002; 2002US-0404809P.
 XX (GETH) GENENTECH INC.
 PA Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
 XX Williams PM, Wu TD, Zhang Z;
 XX WPI; 2003-354551/33.
 DR P-PSDB; ABP97218.
 XX New antibodies against tumor-associated antigenic target polypeptide,
 PT useful for treating or diagnosing tumors or cancers in mammals, e.g.
 PT prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
 PT carcinomas.
 XX Claim 2; Fig 44; 285pp; English.
 XX ACC49493 to ACC49552 encode the human tumour-associated antigenic target
 CC (TAT) proteins given in ABP97175 to ABP97234. The present invention
 CC describes an isolated antibody that binds to a polypeptide having at
 CC least 80 % sequence identity to any of the 60 150-800 residue amino acid
 CC sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
 CC its associated signal peptide, encoded by any of the 60 2000-3000 base
 CC pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
 CC cytosolic activity. The antibody can be used for treating or diagnosing
 CC tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
 CC cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
 CC cell carcinomas, or thyroid cancer.
 XX Sequence 1850 BP; 477 A; 472 C; 498 G; 403 T; 0 U; 0 Other;
 SQ
 Query Match 87.3%; Score 926; DB 7; Length 1850;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 81 GTCTCTCGCAAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG 140
 Db 76 GTCTCTCGCAAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGG 135
 QY 141 TGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
 Db 136 TGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195
 QY 201 CACGTCCTTTCATCAGCGCGCACAGCGCGGCGGAGAGAGAGTGCCTGTCTCTCAG 260
 Db 196 CACGTCCTTTCATCAGCGCGCACAGCGCGGCGGAGAGAGAGTGCCTGTCTCTCAG 255
 QY 261 AAGATGCTGTGGCTCTCGAGAGGCAAGTGTAGGGAACGGAATCCAGAGCGCAGG 320
 Db 256 AAGATGCTGTGGCTCTCGAGAGGCAAGTGTAGGGAACGGAATCCAGAGCGCAGG 315
 QY 321 TCTAGCGCCCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 380
 Db 316 TCTAGCGCCCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 375
 QY 381 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 440
 Db 376 GCTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGCGAGCGAGATCGCTCGGCCA 435

441 CCATCTCGCTGTTCAGACGGGAGAGAGCCGCCACCCCTTACAGGGCCCTTCCACCTCCAGC 500
 436 CCATCTCGCTGTTCAGACGGGAGAGAGCCGCCACCCCTTACAGGGCCCTTCCACCTCCAGC 495
 501 TTCCGGACCCCGAGCAGCAGCTGGAGTGAACCGGGAGTCCGTGGCGGACCCGCAACA 560
 496 TTCCGGACCCCGAGCAGCAGCTGGAGTGAACCGGGAGTCCGTGGCGGACCCGCAACA 555
 561 GAACCATCTTTCAGCAGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 620
 556 GAACCATCTTTCAGCAGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 615
 621 GCAGTAACTTCGGGCAATCAGGCGCACTGTCTAGCGCAGCGGCGGCGCATGAGAGGCGCC 680
 616 GCAGTAACTTCGGGCAATCAGGCGCACTGTCTAGCGCAGCGGCGGCGCATGAGAGGCGCC 675
 681 CGCCCACTTACAGCAGGCTCATGGCCACTTACCGGGGCTCTCTTCCAGGACCCAGCAGA 740
 676 CGCCCACTTACAGCAGGCTCATGGCCACTTACCGGGGCTCTCTTCCAGGACCCAGCAGA 735
 741 GCAGTGGGCGCGGCTCTTCTGTGGAGGAGCGCGGCTCCACCAACACACACATCGCGCC 800
 736 GCAGTGGGCGCGGCTCTTCTGTGGAGGAGCGCGGCTCCACCAACACACACATCGCGCC 795
 801 TAGAGAGCGCAGCCATCTCGAGCAAGAGAGATTAACAGAAAGGACACCTCTCTTAGG 860
 796 TAGAGAGCGCAGCCATCTCGAGCAAGAGAGATTAACAGAAAGGACACCTCTCTTAGG 855
 861 GTCCCGCGGGGCGCGGCTGGGCTGCTAGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 920
 856 GTCCCGCGGGGCGCGGCTGGGCTGCTAGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
 921 TAGAGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 980
 916 TAGAGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
 981 CACCTCCCTGTATATAATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
 976 CACCTCCCTGTATATAATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
 1041 GCTTGCAGAAAAA 1057
 1036 GCTTGCAGAAAAA 1052

RESULT 10

AAD60105
 ID AAD60105 standard; cDNA; 1140 BP.
 XX AC
 XX AAD60105;
 DT 18-DEC-2003 (first entry)
 DE Human androgen-regulated gene (ARG), PMPAL.
 XX Human androgen-regulated gene (ARG); PMPAL; prostate cancer;
 KW Human; androgen-regulated gene; ARG; PMPAL; prostate cancer;
 XX chromosome 20q13; gene; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PH CDS 95..853
 FT /*tag= a
 FT /product= "Human PMPAL protein"
 FT /note= "CDS is referred to as SEQ ID NO:2 in claim 1 of
 FT the specification"
 XX
 XX US6566130-B1.
 XX
 XX 20-MAY-2003.
 XX
 XX 26-JAN-2001; 2001US-00769482.

XX 28-JAN-2000; 2000US-0178772P.
 PR 31-JAN-2000; 2000US-0179045P.
 XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
 XX Srivastava S, Moul JW, Xu LL, Segawa T;
 PI P-PSDB; AAB39589.
 DR WPI; 2003-719644/68.
 XX Novel isolated androgen-regulated gene designated as PMPAL useful for
 PT selecting primers and probes for detecting prostate cancer cells in
 PT biological samples by nucleic acid amplification techniques.
 XX Example 7; Col 29-30; 58pp; English.
 XX The invention relates to an isolated androgen-regulated gene (ARG)
 CC designated as PMPAL. The invention is useful for selecting primers and
 CC probes for detecting prostate cancer cells in a biological sample by
 CC using nucleic acid amplification techniques. The present sequence is
 CC human PMPAL cDNA. PMPAL gene chromosome 20q13
 XX
 SQ Sequence 1140 BP; 271 A; 349 C; 336 G; 184 T; 0 U; 0 Other;
 Query Match 86.0%; Score 912; DB 9; Length 1140;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 95 CAGGCAATCGCGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGT 154
 DB 89 CAGGCAATCGCGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGT 148
 QY 155 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 214
 DB 149 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 208
 QY 215 AGCGGCAACAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 274
 DB 209 AGCGGCAACAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 268
 QY 275 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 334
 DB 269 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 328
 QY 335 CGGCGCACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 394
 DB 329 CGGCGCACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 388
 QY 395 CAGCGCACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 454
 DB 389 CAGCGCACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 448
 QY 455 CAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 514
 DB 449 CAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 508
 QY 515 CAGCGAGTGAACCTGAACCGGGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 574
 DB 509 CAGCGAGTGAACCTGAACCGGGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 568
 QY 575 AGTGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 634
 DB 569 AGTGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 628
 QY 635 ATCAGCGCCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 694
 DB 629 ATCAGCGCCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 688
 QY 695 GAGGTATCGGCGCACTTACCGGGGTTCTCTTTCAGACACAGCAGAGCAGTGTGGGCGGCC 754
 DB 689 GAGGTATCGGCGCACTTACCGGGGTTCTCTTTCAGACACAGCAGAGCAGTGTGGGCGGCC 748

QY 755 TCCTTGCTGAGGAGGACCCGGCTCCACACACACATGCGCCCTTAGAGAGCGAGCC 814
 Db 749 TCCTTGCTGAGGAGGACCCGGCTCCACACACACATGCGCCCTTAGAGAGCGAGCC 808
 QY 815 ATCTGAGCAAGAGGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGC 874
 Db 809 ATCTGAGCAAGAGGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGC 868
 QY 875 CGGCTGGGGCTCGTAGGTGAAAGGAGGAGCACTCCGGCTCTCTTAGAGAGGAGTGA 934
 Db 869 CGGCTGGGGCTCGTAGGTGAAAGGAGGAGCACTCCGGCTCTCTTAGAGAGGAGTGA 928
 QY 935 GAGAGAGCGGGGGCCAGCAAGCATGCTGTGGCCCTCCCTCCCACTCTCTGTGTGA 994
 Db 929 GAGAGAGCGGGGGCCAGCAAGCATGCTGTGGCCCTCCCTCCCACTCTCTGTGTGA 988
 QY 995 TAAATATTACATGATGCTGTCTGTGAATGACACAGCTTAAGAGAGCTTGCAGAAAAA 1054
 Db 989 TAAATATTACATGATGCTGTCTGTGAATGACACAGCTTAAGAGAGCTTGCAGAAAAA 1048
 QY 1055 AAA 1057
 Db 1049 AAA 1051

RESULT 11
 AAA75164
 ID AAA75164 standard; cDNA; 969 BP.
 AC AAA75164;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE cDNA clone encoding a human TANGO 261 polypeptide.
 XX
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
 KW cellular proliferation; cellular differentiation; cellular adhesion;
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
 KW intestinal disorder; spleen associated disease; renal disorder;
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
 KW brain herniation; iatrogenic disease; inflammation; meningitis;
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 6..764
 FT /*tag= a
 FT /product= "TANGO 261"
 XX
 FN WO200052022-A1.
 XX
 XX 08-SEP-2000.
 XX
 XX 01-MAR-2000; 2000WO-US005226.
 XX
 XX 01-MAR-1999; 99US-0122458P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
 XX
 XX WPI; 2000-579269/54.
 XX P-PSDB; AAB18462.
 XX
 XX Novel human and murine secreted proteins designated TANGO 216, 261, 262,
 XX 266 and 267 useful as modulating agents of cellular processes, e.g. for
 XX treating cancer.
 XX
 XX Disclosure; Page; 175pp; English.

XX AAA75163-65 encode human TANGO 261 proteins. The specification also
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
 CC polypeptides can be used to modulate cellular proliferation, modulate
 CC cellular differentiation and/or modulate cellular adhesion. The proteins
 CC can be used to treat any von Willebrand factor-associated disorder,
 CC regulate extracellular matrix structuring, cellular adhesion, and cell
 CC trafficking and/or migration, modulate cellular interactions, modulate the
 CC cell adhesion in proliferative disorders, such as cancer, modulate the
 CC proliferation, differentiation, and/or function of cells that appear in
 CC the bone marrow, and leukocytes, treat bone marrow, blood and
 CC hematopoietic associated diseases and disorders, atelectasis, pulmonary
 CC congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and
 CC bronchiectasis, intestinal disorders, spleen associated diseases,
 CC modulate renal disorders, treat cardiovascular disorders such as ischemic
 CC heart disease, modulate the proliferation, differentiation, and/or
 CC function of bone and cartilage cells and to treat bone and/or cartilage
 CC associated diseases or disorder. They may also be used to treat disorders
 CC associated with the ovaries, and cerebral oedema, hydrocephalus, brain
 CC herniations, iatrogenic disease, inflammation, bacterial and viral
 CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
 CC disease, multiple sclerosis, brain cancers, hydrocephalus and
 CC encephalitis, and treat hepatic disorders. note: the present sequence
 CC does not appear in the specification; it was created using information
 CC provided
 XX
 SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 U; 0 Other;
 Query Match 85.8%; Score 910; DB 3; Length 969;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 960; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 101 ATGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTGATGATGATGATGATG 160
 Db 6 ATGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTGATGATGATGATGATG 65
 QY 161 GTGTGATCATGCTGCTGCTGAGCCTACAGCTGTCTGACCGTCTTATCATGACGG 220
 Db 66 GTGTGATCATGCTGCTGCTGAGCCTACAGCTGTCTGACCGTCTTATCATGACGG 125
 QY 221 CACAGCCAGGCGGAGGAGAGATGCTGCTCCTCAGAGAGATGCTGTGGCCCTCG 280
 Db 126 CACAGCCAGGCGGAGGAGAGATGCTGCTCCTCAGAGAGATGCTGTGGCCCTCG 185
 QY 281 GAGAGCAGTGTGAGGCAACGGAAATCCAGAGCGCGAGGTCTTACGCCGCCCTCGGCC 340
 Db 186 GAGAGCAGTGTGAGGCAACGGAAATCCAGAGCGCGAGGTCTTACGCCGCCCTCGGCC 245
 QY 341 ACCGACCGCTGGCGCTGCGCCCTTCGCGCGGAGCGCTTCCACCGCTTCAGGCC 400
 Db 246 ACCGACCGCTGGCGCTGCGCCCTTCGCGCGGAGCGCTTCCACCGCTTCAGGCC 305
 QY 401 ACCTATCGTACCTGAGCAGCAGATGACCTGCGCGCCACCATCTCGCTGTGACGCGG 460
 Db 306 ACCTATCGTACCTGAGCAGCAGATGACCTGCGCGCCACCATCTCGCTGTGACGCGG 365
 QY 461 GAGGAGCCCCCACCCTTACAGGCGCCCTGACCCCTTCCAGCTTTCGGGACCCCGAGCAG 520
 Db 366 GAGGAGCCCCCACCCTTACAGGCGCCCTGACCCCTTCCAGCTTTCGGGACCCCGAGCAG 425
 QY 521 CTGGAACTGAACCGGAGTGTGCGGCGACCCGCCAAAGAACCATCTTCGACAGTAC 580
 Db 426 CTGGAACTGAACCGGAGTGTGCGGCGACCCGCCAAAGAACCATCTTCGACAGTAC 485
 QY 581 CTGATGATAGTCCAGGCTGGGCGGCCCTTGGCCCCCAGCAGTAATCTGGGCACTACG 640
 Db 486 CTGATGATAGTCCAGGCTGGGCGGCCCTTGGCCCCCAGCAGTAATCTGGGCACTACG 545
 QY 641 GCCACGTGTCTACGCGCGCGCGCATGAGAGGGCGCGGCCACCTTACAGCGAGTTC 700
 Db 546 GCCACGTGTCTACGCGCGCGCGCATGAGAGGGCGCGGCCACCTTACAGCGAGTTC 605
 QY 701 ATCGGCCACTACCCGGGGTCTCTCTTCAGACACAGCAGAGCAGTGGGCCCTCTCTTG 760

Db 606 ATCGGCCACTACCCGGGTCCTCTTCCAGCACCAGCAGCAGTGGCGCGCTCTCTTG 665

Qy 761 CTGAGGGGACCCGGCTCCACACACACACATCGCGCCCTAGAGAGCCAGCATCTGG 820

Db 666 CTGAGGGGACCCGGCTCCACACACACACATCGCGCCCTAGAGAGCCAGCATCTGG 725

Qy 821 AGCAAGAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT 880

Db 726 ACNAAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT 785

Qy 881 GGGGCTCGTAGGTGAAAGGAGAGACATCCCGCGCTCTTAGAGAGAGAGTGAAGGAA 940

Db 786 GGGGCTCGTAGGTGAAAGGAGAGACATCCCGCGCTCTTAGAGAGAGAGTGAAGGAA 845

Qy 941 GGGGGGGGGGAGCAGCAGCAGTGTGGCGCTTCCCTCCACCTCCCTGTCTATTAATA 1000

Db 846 GGGGGGGGGGAGCAGCAGCAGTGTGGCGCTTCCCTCCACCTCCCTGTCTATTAATA 905

Qy 1001 TTATCATGTGATGTCTGTGCTGAATGCACAAAGCTAGAGAGCTTGCAAAAAA 1060

Db 906 TTATCATGTGATGTCTGTGCTGAATGCACAAAGCTAGAGAGCTTGCAAAAAA 965

Qy 1061 A 1061

Db 966 A 966

RESULT 12

AAA75165 standard; cdNA; 969 BP.

XX AAA75165;

XX 15-JAN-2001 (first entry)

XX cDNA clone encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;

XX cellular proliferation; cellular differentiation; cellular adhesion;

XX von Willebrand factor-associated disorder; cell trafficking; cancer;

XX hematopoietic associated disease; atelectasis; pulmonary congestion;

XX oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;

XX intestinal disorder; spleen associated disease; renal disorder;

XX cardiovascular disorder; ischemic heart disease; hydrocephalus;

XX brain herniation; iatrogenic disease; inflammation; meningitis;

XX Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;

XX multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 6..764

XX /*tag= a

XX /product= "TANGO 261"

XX WO200052022-A1.

XX 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US005226.

XX 01-MAR-1999; 99US-0122458P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX WPI: 2000-579269/54.

XX P-PSDB; AAB18463.

XX Novel human and murine secreted proteins designated TANGO 216, 261, 262,

XX 266 and 267 useful as modulating agents of cellular processes, e.g. for

PT treating cancer.

XX Disclosure; Page; 175pp; English.

XX AAA75163-65 encode human TANGO 261 proteins. The specification also

XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO

XX polypeptides can be used to modulate cellular proliferation, modulate

XX cellular differentiation and/or modulate cellular adhesion. The proteins

XX can be used to treat any von Willebrand factor-associated disorder,

XX regulate extracellular matrix structuring, cellular adhesion, and cell

XX trafficking and/or migration, modulate cellular interactions, modulate

XX cell adhesion in proliferative disorders, such as cancer, modulate the

XX proliferation, differentiation, and/or function of cells that appear in

XX the bone marrow, and leukocytes, treat bone marrow, blood and

XX hematopoietic associated diseases and disorders, atelectasis, pulmonary

XX congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and

XX bronchiectasis, intestinal disorders, spleen associated diseases,

XX modulate renal disorders, treat cardiovascular disorders such as ischemic

XX heart disease, modulate the proliferation, differentiation, and/or

XX function of bone and cartilage cells and to treat bone and/or cartilage

XX associated diseases or disorder. They may also be used to treat disorders

XX associated with the ovaries, and cerebral oedema, hydrocephalus, brain

XX herniations, iatrogenic disease, inflammations, bacterial and viral

XX meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's

XX disease, multiple sclerosis, brain cancers, hydrocephalus and

XX encephalitis, and treat hepatic disorders, note: the present sequence

XX does not appear in the specification; it was created using information

XX provided

XX Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 U; 0 Other;

Qy Query Match 85.8%; Score 910; DB 3; Length 969;

Db Best Local Similarity 99.9%; Pred. No. 0;

Qy Matches 960; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 101 ATGCGGAGCTGGAGTTTCTTTCAGATCATCATCATCTGCTGATGATGATGATGATG 160

Db 6 ATGCGGAGCTGGAGTTTCTTTCAGATCATCATCATCTGCTGATGATGATGATGATG 65

Qy 161 GTGTGTATCATCATCTGCTGCTGAGCCACTACAGCTCTCTGACGGTCTTTCATCAGCCGG 220

Db 66 GTGTGTATCATCATCTGCTGCTGAGCCACTACAGCTCTCTGACGGTCTTTCATCAGCCGG 125

Qy 221 CACAGCCAGGGCGGAGGAGAGAGATGCTCTGTCTCAGAGAGATGCTCTGCTGCTG 280

Db 126 CACAGCCAGGGCGGAGGAGAGAGATGCTCTGTCTCAGAGAGATGCTCTGCTGCTG 185

Qy 281 GAGAGCAGATGTCTAGGCAACCGAATCCAGAGCCGAGGTCTACGCCCGCGCTCGGCC 340

Db 186 GAGAGCAGATGTCTAGGCAACCGAATCCAGAGCCGAGGTCTACGCCCGCGCTCGGCC 245

Qy 341 ACCGAGCCGCTGCGCGCTGCGCGCTTTCGCGCCAGCGGGAGCGCTTTCACCGCTTTCAGCC 400

Db 246 ACCGAGCCGCTGCGCGCTGCGCGCTTTCGCGCCAGCGGGAGCGCTTTCACCGCTTTCAGCC 305

Qy 401 ACCTATTCGCTACCTGCGAGCAACAGATCGACTGCGCGCCCAACATCTGCTGTGAGAGCGG 460

Db 306 ACCTATTCGCTACCTGCGAGCAACAGATCGACTGCGCGCCCAACATCTGCTGTGAGAGCGG 365

Qy 461 GAGGAGCCCGGACCGCTTACCGAGGGCGCTGACCGCTTTCAGCTTTCGCGAGCCGAGCAG 520

Db 366 GAGGAGCCCGGACCGCTTACCGAGGGCGCTTTCAGCTTTCGCGAGCCGAGCAGCAG 425

Qy 521 CTGGAACTGAACCGGAGTGTGCGCGCAACCGCCCAACAGAACACCTTTCGAGCAGTGAC 580

Db 426 CTGGAACTGAACCGGAGTGTGCGCGCAACCGCCCAACAGAACACCTTTCGAGCAGTGAC 485

Qy 581 CTGATGATGATGCTCAGAGCTGGCGCGCTGCGCGCCCGCGCGCGCGCGCGCGCGCGCG 640

Db 486 CTGATGATGATGCTCAGAGCTGGCGCGCTGCGCGCCCGCGCGCGCGCGCGCGCGCGCG 545

Qy 641 GCCACGTGTCTAGCGCAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700

RESULT 15	
ACC49537	
ID	ACC49537 standard; cDNA; 806 BP.
XX	
AC	ACC49537;
AC	
XX	
DT	01-JUL-2003 (first entry)
XX	
DE	Tumour-associated antigenic target protein TAT375 cDNA SEQ ID NO:45.
XX	
KW	Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
KW	cancer; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO2003024392-A2.
XX	
PD	27-MAR-2003.
XX	
PF	11-SEP-2002; 2002WO-US028859.
XX	
PR	18-SEP-2001; 2001US-033268P.
PR	19-OCT-2001; 2001US-0339227P.
PR	07-NOV-2001; 2001US-0336827P.
PR	20-NOV-2001; 2001US-0331906P.
PR	02-JAN-2002; 2002US-034544P.
PR	03-APR-2002; 2002US-0369724P.
PR	19-AUG-2002; 2002US-0404809P.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
PI	Williams PM, Wu TD, Zhang Z;
XX	
WI	WFI; 2003-354551/33.
DR	P-PSDB; ABP97219.
XX	
PT	New antibodies against tumor-associated antigenic target polypeptide
PT	useful for treating or diagnosing tumors or cancers in mammals, e.g.
PT	prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
PT	carcinomas.
XX	
PS	Claim 2; Fig 45; 285pp; English.
XX	
CC	ACC49493 to ACC49552 encode the human tumour-associated antigenic ta
CC	(TAT) proteins given in ABP97175 to ABP97234. The present invention
CC	describes an isolated antibody that binds to a polypeptide having at
CC	least 80 % sequence identity to any of the 60 150-800 residue amino
CC	sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lack
CC	its associated signal peptide, encoded by any of the 60 2000-3000 ba
CC	pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins hav
CC	cytostatic activity. The antibody can be used for treating or diagno
CC	tumours or cancers in mammals, e.g. prostate cancer, lung cancer, br
CC	cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, ren
CC	cell carcinomas, or thyroid cancer
XX	
SQ	Sequence 806 BP; 161 A; 286 C; 241 G; 118 T; 0 U; 0 Other;
	Query Match 74.3%; Score 788; DB 7; Length 806;
	Best Local Similarity 100.0%; Pred. No. 2.1e-304;
	Matches 788, Conservative 0; Mismatches 0; Indels 0; Gaps
Qy	81 GTCTCTGCCAAACCCAGGCATATCGCGAGCTGGAGTTGTTCAGATCATCATCATCGT
Db	13 GTCTCTCTGCCAAACCCAGGCATATCGCGAGCTGGAGTTGTTCAGATCATCATCATCGT
Qy	141 TGGTGATGATGGTCAATGCTGTGTGTGTATCAGTCTGCTGAGCCACTACAGCTGTCT
Db	73 TGGTGATGATGGTCAATGCTGTGTGTGTATCAGTCTGCTGAGCCACTACAGCTGTCT
Qy	201 CACGGTCTCTTCATCAGCCGGCACAAGCCAGCGGCGGAGAGAGATGCCCTGTCTCTCT

; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-5

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 GGCAGAACACTCCGCGCTTCTTAG 923
|||||
DB 1 GGCAGAACACTCCGCGCTTCTTAG 24

RESULT 4
US-09-769-482-6/c
; Sequence 6: Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-6

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1022 GAATGCACAACTAAGAGAGCTTG 1045
|||||
DB 24 GAATGCACAACTAAGAGAGCTTG 1

RESULT 5
US-09-769-482-10/c
; Sequence 10: Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-10

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1022 GAATGCACAACTAAGAGAGCTTG 1045
|||||
DB 24 GAATGCACAACTAAGAGAGCTTG 1

RESULT 6
US-08-332-766A-1
; Sequence 1: Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-1

Query Match 2.2%; Score 23; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGTGGTGTGATGGTGTGATGGTG 160
DB 123 TGTGGTGTGATGGTGTGATGGTG 145

RESULT 7

US-09-769-482-7
; Sequence 7, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-7

Query Match 2.0%; Score 21; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTGGGTTCTGGTGAAGGC 26
DB 1 CTGGGTTCTGGTGAAGGC 21

RESULT 8

US-08-522-421-7
; Sequence 7, Application US/08522421
; Patent No. 5908973
; GENERAL INFORMATION:
; APPLICANT: Abu-Bakar, Umi Kalsom
; APPLICANT: Barton, Sarah Louise
; APPLICANT: Gallego-Veigas, Pedro Pablo
; APPLICANT: Gray, Julie Elizabeth
; APPLICANT: Grierson, Donald
; APPLICANT: Lowe, Alexandra Louise
; APPLICANT: Picton, Steve
; APPLICANT: Whotton, Lee Colin
; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS
; DERIVED THEREFROM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/522,421
; FILING DATE: 11-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305868.3
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305869.1
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305859.2
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305865.9
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305866.7
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305867.5
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305860.0
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305862.6
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9314351.9
; FILING DATE: 12-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9320988.0
; FILING DATE: 12-OCT-1993
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: BRT17
US-08-522-421-7

Query Match 2.0%; Score 21; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1041 GCTTGCACAAAAA 1061
DB 663 GCTTGCACAAAAA 683

RESULT 9

US-08-165-315D-3
; Sequence 3, Application US/08165315D
; Patent No. 5525716
; GENERAL INFORMATION:
; APPLICANT: Odd-Arne Olsen
; APPLICANT: Roger Kalla
; TITLE OF INVENTION: Promoter
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: St. Onge, Steward, Johnston & Reens
; STREET: 986 Bedford Street
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06905
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/165,315D
FILING DATE: 10 December 1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9324707.0
FILING DATE: 2 December 1993
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: 2105-P0001
TELEPHONE: 201-324-6155
TELEFAX: 201-327-1096
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
DESCRIPTION: gene
FRAGMENT TYPE: gene
FEATURE:
NAME/KEY: Ltp2 gene
US-08-165-315D-3

Query Match 2.0%; Score 21; DB 1; Length 1327;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GTGATGATGGTGGTGGTG 163
DB 918 GTGATGATGGTGGTGGTG 938

RESULT 10
US-08-046-585-15/c
Sequence 15, Application US/08046585
Patent No. 5453362
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBRACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/046,585
FILING DATE: 12-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57503-1/RAO
TELEPHONE: (415) 781-1989
TELEFAX: 910 277299

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-046-585-15

Query Match 2.0%; Score 21; DB 1; Length 8252;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGGTGGTGATGG 158
DB 2748 TGGTGGTGATGGTGGTGATGG 2728

RESULT 11
US-08-393-703-15/c
Sequence 15, Application US/08393703
Patent No. 5585239
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBRACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,703
FILING DATE: 24-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57503-2/RAO
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-393-703-15

Query Match 2.0%; Score 21; DB 1; Length 8252;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGGTGGTGATGG 158
DB 2748 TGGTGGTGATGGTGGTGATGG 2728

RESULT 12
PCT-US93-11721-15/c

Sequence 15, Application PC/TUS9311721
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBRACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11721
FILING DATE: 03-DEC-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: PP-57503-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US93-11721-15

Query Match 2.0%; Score 21; DB 5; Length 8252;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGATGGTGATGG 158
|||||

Db 2748 TGGTGGTGATGATGGTGATGG 2728

RESULT 13
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g

```
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```

```
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

Query Match          2.0%; Score 21; DB 4; Length 1664976;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 TGGTGGTGATGATGGTGGTGG 158
      |||||
Db 568047 TGGTGGTGATGATGGTGGTGG 568027

RESULT 14
US-09-769-482-9
; Sequence 9, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
; US-09-769-482-9

Query Match          1.9%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 363 CCTTGCCCGAGCGGAGCGC 382
      |||||
Db 1 CCTTGCCCGAGCGGAGCGC 20

RESULT 15
US-08-753-247-22/c
; Sequence 22, Application US/08753247
; Patent No. 6210929
; GENERAL INFORMATION:
; APPLICANT: SCHLOKAT, Uwe
; APPLICANT: FISCHER, Bernhard
; APPLICANT: FALKNER, Falko-Guenther
; APPLICANT: DORNER, Friedrich
; APPLICANT: EIBL, Johann
; TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A PURIN
; TITLE OF INVENTION: DERIVATIVE OR A DERIVATIVE OF A PURIN ANALOGUE AND A
; TITLE OF INVENTION: HETEROLOGOUS SEQUENCE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
```

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/753,247
 FILING DATE: 22-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: AT 1928/95
 FILING DATE: 24-NOV-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 40433/149
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 50 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-753-247-22

Query Match 1.9% Score 20; DB 3; Length 50;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 TGATGATGGTGGTGGTG 163
 |||||
 Db 42 TGATGATGGTGGTGGTG 23

Search completed: May 26, 2004, 20:27:26
 Job time : 113 secs


```

1 PRIOR FILING DATE: 1999-12-23
2 PRIOR APPLICATION NUMBER: 09/474,071
3 PRIOR FILING DATE: 1999-12-29
4 PRIOR APPLICATION NUMBER: 09/474,072
5 PRIOR FILING DATE: 1999-12-29
6 PRIOR APPLICATION NUMBER: 09/514,010
7 PRIOR FILING DATE: 2000-02-25
8 PRIOR APPLICATION NUMBER: 09/516,745
9 PRIOR FILING DATE: 2000-03-01
10 PRIOR APPLICATION NUMBER: 09/572,002
11 PRIOR FILING DATE: 2000-05-14
12 PRIOR APPLICATION NUMBER: 09/597,993
13 PRIOR FILING DATE: 2000-06-19
14 PRIOR APPLICATION NUMBER: 09/599,596
15 PRIOR FILING DATE: 2000-06-22
16 PRIOR APPLICATION NUMBER: 09/630,334
17 PRIOR FILING DATE: 2000-07-31
18 PRIOR APPLICATION NUMBER: 09/606,565
19 PRIOR FILING DATE: 2000-06-29
20 PRIOR APPLICATION NUMBER: 09/606,317
21 PRIOR FILING DATE: 2000-06-29
22 PRIOR APPLICATION NUMBER: 09/665,666
23 PRIOR FILING DATE: 2000-09-20
24 PRIOR APPLICATION NUMBER: 09/677,751
25 PRIOR FILING DATE: 2000-09-30
26 NUMBER OF SEQ ID NOS: 162
27 SEQ ID NO 55
28 LENGTH: 969
29 TYPE: DNA
30 ORGANISM: Homo sapiens
31 FEATURE:
32 NAME/KEY: CDS
33 LOCATION: (6)...(761)
34 US-09-796-753-55

Query Match          90.6%; Score 961; DB 10; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGCGAGCTGAGTTGTTTCAGATCATCATCATGTTGTTGATGATGTTGATGTTG 160
DB 6 ATGGCGAGCTGAGTTGTTTCAGATCATCATCATGTTGTTGATGATGTTGATGTTG 65

QY 161 GTGGTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 220
DB 66 GTGGTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125

QY 221 CACAGCCAGGGCGGAGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
DB 126 CACAGCCAGGGCGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185

QY 281 GAGAGCAGATGTCAGGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
DB 186 GAGAGCAGATGTCAGGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245

QY 341 ACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
DB 246 ACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305

QY 401 ACTATCCGTTCTGAGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
DB 306 ACTATCCGTTCTGAGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365

QY 461 GAGGAGCCCGCTTACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
DB 366 GAGGAGCCCGCTTACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425

QY 521 CTGGAAGTAACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 580
DB 426 CTGGAAGTAACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 485

QY 581 CTGATGATAGTGCAGGCTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640

```

```

DB 486 CTGATGATAGTGCAGGCTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
QY 641 GCCAGTGTCTACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700
DB 546 GCCAGTGTCTACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 605
QY 701 ATCGGCCACTACCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
DB 606 ATCGGCCACTACCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 665
QY 761 CTGAGGCGGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 820
DB 666 CTGAGGCGGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 725
QY 821 AGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 880
DB 726 AGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 785
QY 881 GGGGCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
DB 786 GGGGCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845
QY 941 GGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1000
DB 846 GGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 905
QY 1001 TTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1060
DB 906 TTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
QY 1061 A 1061
DB 966 A 966

RESULT 2
US-10-205-823-412
; Sequence 412, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-10-205-823-412

```

Query Match      87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 81 GTCTCTCGGAAACAGGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 140
Db 76 GTCTCTCGGAAACAGGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 135

Qy 141 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
Db 136 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195

Qy 201 CACGGTCTTTCATCAGCGCGGACAGGAGGCGGAGGAGAGAGATGCTGCTCTCAG 260
Db 196 CACGGTCTTTCATCAGCGCGGACAGGAGGCGGAGGAGAGATGCTGCTCTCAG 255

Qy 261 AAGGATGCTGTGGCCCTTCGAGAGGACAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 320
Db 256 AAGGATGCTGTGGCCCTTCGAGAGGACAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 315

Qy 321 TCTAGGCGCGGCTTCGCGGCGGACCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 380
Db 316 TCTAGGCGCGGCTTCGCGGCGGACCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 375

Qy 381 GCTTCCACCGCTTCAGCGGCGGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 440
Db 376 GCTTCCACCGCTTCAGCGGCGGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 435

Qy 441 CCATCTCGCTGTTCAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 500
Db 436 CCATCTCGCTGTTCAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 495

Qy 501 TTGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
Db 496 TTGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555

Qy 561 GAACCATCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 620
Db 556 GAACCATCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 615

Qy 621 GCAGTAATCTCGGCGCATCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 680
Db 616 GCAGTAATCTCGGCGCATCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 675

Qy 681 GCGCCCATCTCAGCGAGGAGTCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 740
Db 676 GCGCCCATCTCAGCGAGGAGTCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 735

Qy 741 GCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 800
Db 736 GCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 795

Qy 801 TAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 860
Db 796 TAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 855

Qy 861 GTCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 920
Db 856 GTCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 915

Qy 921 TAGAAGAGGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 980
Db 916 TAGAAGAGGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 975

Qy 981 CACCTCCCTGTGTATAAATTTTATCATGTGATGTCTGTGTGTGATGTGATGTGATGTG 1040
Db 976 CACCTCCCTGTGTATAAATTTTATCATGTGATGTCTGTGTGTGATGTGATGTGATGTG 1035

Qy 1041 GCTTGCAAAAAA 1057
Db 1036 GCTTGCAAAAAA 1052

```

```

RESULT 3
US-10-822-208
; Sequence 208, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Subhang
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029F2RNM
; CURRENT APPLICATION NUMBER: US/10/301.822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(854)
US-10-301-822-208

```

```

Query Match      87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 81 GTCTCTCGGAAACAGGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 140
Db 76 GTCTCTCGGAAACAGGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 135

Qy 141 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
Db 136 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195

Qy 201 CACGGTCTTTCATCAGCGCGGACAGGAGGCGGAGGAGAGATGCTGCTCTCAG 260
Db 196 CACGGTCTTTCATCAGCGCGGACAGGAGGCGGAGGAGAGATGCTGCTCTCAG 255

Qy 261 AAGGATGCTGTGGCCCTTCGAGAGGACAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 320
Db 256 AAGGATGCTGTGGCCCTTCGAGAGGACAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 315

Qy 321 TCTAGGCGCGGCTTCGCGGCGGACCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 380
Db 316 TCTAGGCGCGGCTTCGCGGCGGACCGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 375

Qy 381 GCTTCCACCGCTTCAGCGGCGGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 440
Db 376 GCTTCCACCGCTTCAGCGGCGGACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 435

Qy 441 CCATCTCGCTGTTCAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 500
Db 436 CCATCTCGCTGTTCAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 495

Qy 501 TTGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
Db 496 TTGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 555

```

QY 561 GAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCA 620
 DB |||||
 QY 556 GAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCA 615
 DB |||||
 QY 621 GCAGTAACCTCGGCGCATCAGCGCCACGCTGCTACCGGCGGCGCGCATGAGGCGCGC 680
 DB |||||
 QY 616 GCAGTAACCTCGGCGCATCAGCGCCACGCTGCTACCGGCGGCGCGCATGAGGCGCGC 675
 DB |||||
 QY 681 CGCCCACTTACAGCGAGTCTACGCGCCACTACCGGCGGCTCTCTTCCAGCACGAGCA 740
 DB |||||
 QY 676 CGCCCACTTACAGCGAGTCTACGCGCCACTACCGGCGGCTCTCTTCCAGCACGAGCA 735
 DB |||||
 QY 741 GCAGTGGCGCGCCCTCTCTTCTGAGGCGGACCGGCTCCACCAACACACATCGCGCCC 800
 DB |||||
 QY 736 GCAGTGGCGCGCCCTCTCTTCTGAGGCGGACCGGCTCCACCAACACACATCGCGCCC 795
 DB |||||
 QY 801 TAGAGAGCGCAGCCATCTGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGG 860
 DB |||||
 QY 796 TAGAGAGCGCAGCCATCTGAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGG 855
 DB |||||
 QY 861 GTCCCGCAGGCGGCGCGGCTGGGCTGCTAGGTGAAAGGCGACACCTCGCGCTTCT 920
 DB |||||
 QY 856 GTCCCGCAGGCGGCGCGGCTGGGCTGCTAGGTGAAAGGCGACACCTCGCGCTTCT 915
 DB |||||
 QY 921 TAGAGAGAGTGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 980
 DB |||||
 QY 916 TAGAGAGAGTGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 975
 DB |||||
 QY 981 CACCTCCCTGTGTATTAATTTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGA 1040
 DB |||||
 QY 976 CACCTCCCTGTGTATTAATTTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGA 1035
 DB |||||
 QY 1041 GCTTGCAGAAAAA 1057
 DB |||||
 QY 1036 GCTTGCAGAAAAA 1052
 DB |||||

RESULT 4

US-10-241-220-44
 ; Sequence 44, Application US/10241220
 ; Publication No. US20030148408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frantz, Gretchen
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Phillips, Heidi
 ; APPLICANT: Polakis, Paul
 ; APPLICANT: Spencer, Susan
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wu, Thomas
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 ; TREATMENT OF TUMOR
 ; FILE REFERENCE: P5010R1-US
 ; CURRENT APPLICATION NUMBER: US/10/241,220
 ; CURRENT FILING DATE: 2002-12-13
 ; NUMBER OF SEQ ID NOS: 120
 ; SEQ ID NO 44
 ; LENGTH: 1850
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-241-220-44

Query Match 87.3%; Score 926; DB 15; Length 1850;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 81 GTCTCTCGCAACAGCGCAATCGGAGCTGGAGTTTGTTCAGATCATCATCATCTGG 140
 DB GTCTCTCGCAACAGCGCAATCGGAGCTGGAGTTTGTTCAGATCATCATCATCTGG 135
 QY 141 TGGTATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
 |||||

DB 136 TGGTATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 195
 QY |||||
 DB 201 CAGGCTCTTTCATCAGCGCCACAGCGGCGGAGAGAGAGATGCGCTGCTCTCAG 260
 QY |||||
 DB 196 CAGGCTCTTTCATCAGCGCCACAGCGGCGGAGAGAGAGATGCGCTGCTCTCAG 255
 QY |||||
 DB 261 AAGGATGCTGTGCGCTTCGAGAGACAGTGTGAGGAAACGGAATCCAGAGCGGAGG 320
 DB |||||
 DB 256 AAGGATGCTGTGCGCTTCGAGAGACAGTGTGAGGAAACGGAATCCAGAGCGGAGG 315
 QY |||||
 DB 321 TCTACGCGCGCTTCGCGCCACAGCGGCTGCGCGCTGCGCGCTGCGCGCGGAGC 380
 DB |||||
 DB 316 TCTACGCGCGCTTCGCGCCACAGCGGCTGCGCGCTGCGCGCTGCGCGCGGAGC 375
 QY |||||
 DB 381 GCTTTCACCGCTTTCAGCGCCACCTATCCGTACCTGACAGCAGAGATGACCTGCGCGCA 440
 DB |||||
 DB 376 GCTTTCACCGCTTTCAGCGCCACCTATCCGTACCTGACAGCAGAGATGACCTGCGCGCA 435
 QY |||||
 DB 441 CCATCTGCTGTGAGAGCGGAGGAGCCCACTTACCGAGGCGCTTGTGACCTCTCAGC 500
 DB |||||
 DB 436 CCATCTGCTGTGAGAGCGGAGGAGCCCACTTACCGAGGCGCTTGTGACCTCTCAGC 495
 QY |||||
 DB 501 TTGCGGACCCGAGCAGCTGGAACCTGAAACGCGGAGTGGTGGCGCACCCCAACA 560
 DB |||||
 DB 496 TTGCGGACCCGAGCAGCTGGAACCTGAAACGCGGAGTGGTGGCGCACCCCAACA 555
 QY |||||
 DB 561 GAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTGCGCCCCCA 620
 DB |||||
 DB 556 GAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTGCGCCCCCA 615
 QY |||||
 DB 621 GCAGTAACCTCGGCGCATCAGCGCCACGCTGCTACCGGCGGCGGCGGCGGAGGCGCGC 680
 DB |||||
 DB 616 GCAGTAACCTCGGCGCATCAGCGCCACGCTGCTACCGGCGGCGGCGGCGGAGGCGCGC 675
 QY |||||
 DB 681 CGCCCACTTACAGCGAGTCTACGCGCCACTACCGGCGGCTCTCTTCCAGCACGAGCA 740
 DB |||||
 DB 676 CGCCCACTTACAGCGAGTCTACGCGCCACTACCGGCGGCTCTCTTCCAGCACGAGCA 735
 QY |||||
 DB 741 GCAGTGGCGCGCCCTCTCTTCTGAGGCGGACCGGCTCCACCAACACACATCGCGCCC 800
 DB |||||
 DB 736 GCAGTGGCGCGCCCTCTCTTCTGAGGCGGACCGGCTCCACCAACACACATCGCGCCC 795
 QY |||||
 DB 801 TAGAGAGCGCAGCCATCTGAGCAAGAGAGATAAACAGAAAGGACACCTCTCTAGG 860
 DB |||||
 DB 796 TAGAGAGCGCAGCCATCTGAGCAAGAGAGATAAACAGAAAGGACACCTCTCTAGG 855
 QY |||||
 DB 861 GTCCCGCAGGCGGCGGCGGCTGGGCTGCTAGGTGAAAGGCGACACCTCGCGCTTCT 920
 DB |||||
 DB 856 GTCCCGCAGGCGGCGGCGGCTGGGCTGCTAGGTGAAAGGCGACACCTCGCGCTTCT 915
 QY |||||
 DB 921 TAGAGAGAGTGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 980
 DB |||||
 DB 916 TAGAGAGAGTGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 975
 QY |||||
 DB 981 CACCTCCCTGTGTATTAATTTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGA 1040
 DB |||||
 DB 976 CACCTCCCTGTGTATTAATTTACATGTGATGTCTGGTCTGAATGCAAGCTAAGAGA 1035
 DB |||||
 QY 1041 GCTTGCAGAAAAA 1057
 DB |||||
 DB 1036 GCTTGCAGAAAAA 1052
 DB |||||

RESULT 5

US-09-821-812-2
 ; Sequence 2, Application US/09821812
 ; Publication No. US20030166520A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lin, Biaoyang
 ; TITLE OF INVENTION: Androgen Regulated Prostate Specific
 ; Nucleic Acids
 ; FILE REFERENCE: P-18 4373
 ; CURRENT APPLICATION NUMBER: US/09/821,812

CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 4527
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (96)...(851)
US-09-821-812-2

Query Match 87.3%; Score 926; DB 10; Length 4527;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 81 GTCTCTCGGAACACGAGCAATGCGGAGCTGGAGTTTGTTCAGATCATCATCATGCTGG 140
DB 76 GTCTCTCGGAACACGAGCAATGCGGAGCTGGAGTTTGTTCAGATCATCATCATGCTGG 135
QY 141 TGGTGTATGATGGTGTGTGTGATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 200
DB 136 TGGTGTATGATGGTGTGTGTGATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 195
QY 201 CACGGTCTCTTCATCAGCGGCACAGCCAGCGGCGGAGAGAGAGATGCCCTGTCTCTCAG 260
DB 196 CACGGTCTCTTCATCAGCGGCACAGCCAGCGGCGGAGAGAGATGCCCTGTCTCTCAG 255
QY 261 AAGGATGCTGTGGCTCTCGAGAGCAAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 320
DB 256 AAGGATGCTGTGGCTCTCGAGAGCAAGTGTTCAGGCAACGGAATCCAGAGCGGAGG 315
QY 321 TCTACGCCCGCGCTGCGCCACACGCGCTGCGCGCTGCGCGCTGCGCGCGCGCGGAGC 380
DB 316 TCTACGCCCGCGCTGCGCCACACGCGCTGCGCGCTGCGCGCTGCGCGCGCGGAGC 375
QY 381 GCTTCACACGCTTCACAGCCACCTATCCGTACCTGACGACGAGATGACCTGCGGCCA 440
DB 376 GCTTCACACGCTTCACAGCCACCTATCCGTACCTGACGACGAGATGACCTGCGGCCA 435
QY 441 CATCTCGCTGTACAGCGGAGAGAGCCGCCACCTACAGAGGCGCTGACACCTTCAGC 500
DB 436 CATCTCGCTGTACAGCGGAGAGAGCCGCCACCTACAGAGGCGCTGACACCTTCAGC 495
QY 501 TTCCGACACCCGAGCAGCAGCTGMACTGAACCGGAGTGTGCGCGACCCGCCAACA 560
DB 496 TTCCGACACCCGAGCAGCAGCTGMACTGAACCGGAGTGTGCGCGACCCGCCAACA 555
QY 561 GAACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGCGCGCGCTGCGCGCGCC 620
DB 556 GAACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGCGCGCGCTGCGCGCGCC 615
QY 621 GCAGTAACTCGGGGATCAGCGCCACGCTGTACGCGAGCGGCGGCGGATGAGAGGCGCG 680
DB 616 GCAGTAACTCGGGGATCAGCGCCACGCTGTACGCGAGCGGCGGCGGATGAGAGGCGCG 675
QY 681 CGCCCACTTACAGAGGTTCATCGGCCACTACCGCGGCTCTCTTCCAGCACACAGAGA 740
DB 676 CGCCCACTTACAGAGGTTCATCGGCCACTACCGCGGCTCTCTTCCAGCACACAGAGA 735
QY 741 GCAGTGGCGCGCTCTCTTCTGAGAGGAGAGCCCGGCTTCCACCAACACATCGCGGCC 800
DB 736 GCAGTGGCGCGCTCTCTTCTGAGAGGAGAGCCCGGCTTCCACCAACACATCGCGGCC 795
QY 801 TAGAGAGCGGAGCATCTGAGCAAGAGAGAGGATATAACAGAAAGACACCTCTCTAGG 860
DB 796 TAGAGAGCGGAGCATCTGAGCAAGAGAGAGGATATAACAGAAAGACACCTCTCTAGG 855
QY 861 GTCCCAAGGCGGCGCGGCTGCGGTGTGAGTGAAGAGGAGAAACACTCTCGCGCTTCT 920
DB 856 GTCCCAAGGCGGCGGCGGCTGCGGTGTGAGTGAAGAGGAGAAACACTCTCGCGCTTCT 915
QY 921 TAGAAGAGGAGTGAAGAGGAGGCGGCGGCGGAGCAACGATCTGTGTGGCCCTCCCTCC 980
```

```
DB 916 TAGAAGAGGAGTGAAGAGGAGGCGGCGGCGGAGCAACGATCTGTGTGGCCCTCCCTCC 975
QY 981 CACTCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGATGCAACAGCTTAAGAGA 1040
DB 976 CACTCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGATGCAACAGCTTAAGAGA 1035
QY 1041 GCTTGCACAAAAA 1057
DB 1036 GCTTGCACAAAAA 1052
```

RESULT 6

US-10-390-045-1
Sequence 1, Application US/10390045
Publication No. US20030170713A1
GENERAL INFORMATION:
APPLICANT: SRIVASTAVA, SHIV
APPLICANT: MOUL, JUDD W.
APPLICANT: XU, LINDA L.
APPLICANT: SEGAWA, TAKESHIKO
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
FILE REFERENCE: 04995.0057-00000
CURRENT FILING DATE: 2003-03-18
PRIORITY APPLICATION NUMBER: US/09/769,482
PRIORITY FILING DATE: 2001-01-26
PRIORITY APPLICATION NUMBER: 60/178,772
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: 60/179,045
PRIORITY FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1140
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(850)
US-10-390-045-1

Query Match 86.0%; Score 912; DB 15; Length 1140;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 95 CAGCAATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCAT 154
DB 89 CAGCAATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCAT 148
QY 155 ATGGTGGTGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGCAAGGTCCTTTCATC 214
DB 149 ATGGTGGTGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGCAAGGTCCTTTCATC 208
QY 215 AGCCGCGCAGCAGCCAGGCGGCGGAGAGAGATGCCCTGTCTCTCAGAGGATGCCCTGTGG 274
DB 209 AGCCGCGCAGCAGCCAGGCGGCGGAGAGAGATGCCCTGTCTCTCAGAGGATGCCCTGTGG 268
QY 275 CCTTGGAGAGCAGTGTTCAGGCAACGGAATCCAGAGCCGAGGTCATCGCCCGGCT 334
DB 269 CCTTGGAGAGCAGTGTTCAGGCAACGGAATCCAGAGCCGAGGTCATCGCCCGGCT 328
QY 335 CGGCCCAACCGACCGCTGCGCGCTTCCGCGCGCTTCCGCGCGGAGCGCTTCCACCGCTTC 394
DB 329 CGGCCCAACCGACCGCTGCGCGCTTCCGCGCGCTTCCGCGCGGAGCGCTTCCACCGCTTC 388
QY 395 CAGCCCACTTATCGTACCTGACGACGAGATCGACCTGCGCGCCACCATCTCGCTGTCA 454
DB 389 CAGCCCACTTATCGTACCTGACGACGAGATCGACCTGCGCGCCACCATCTCGCTGTCA 448
QY 455 GACGGGAGGAGGCGCCCAACCTTACAGGCGCTTCCGCGCGCTTCCGCGCGGAGCGCCGAG 514
```


730	ATCGTACCTGTGAGCA	CGAGATGACACTGACCACCA	CCATCTCGCTCTCAGACGGGAGG	789
465	AGCCCCACCCCTAC	CAGGGCCCCCTGACAC	CCCTCCACGCTTCGGGAGCCCGCAGCAGCAGCTGG	524
790	AGCCCCACCCCTAC	CAGGGCCCCCTGACAC	CCCTCCACGCTTCGGGAGCCCGCAGCAGCAGCTGG	849
525	AACCTGAACCGGGAGT	CGGTGGCGGACCCGCCAAACAGAAC	CCATCTTCGACAGGTGACCTGA	584
850	AACCTGAACCGGGAGT	CGGTGGCGGACCCGCCAAACAGAAC	CCATCTTCGACAGGTGACCTGA	909
585	TGGATAGTGC	CAGCGTCGGCGGGCCCCCTG	CCCCCCCACGAGTAACCTCTCGGGCATCAGCGCCA	644
910	TGGATAGTGC	CAGCGTCGGCGGGCCCCCTG	CCCCCCCACGAGTAACCTCTCGGGCATCAGCGCCA	969
645	CGTGCTACGGCAG	CGGGGGCGCATGGAGGGCGCGCCGCCAC	CTACAGCAGGTCAATCG	704
970	CGTGCTACGGCAG	CGGGGGCGCATGGAGGGCGCGCCGCCAC	CTACAGCAGGTCAATCG	1029
705	GCCTACTACCCGGGT	CTCTCTTCAGCAGC	CAGCAGCAGCTGGGCGCGCCCTCTTGGCTGG	764
1030	GCCACTACCCGGGT	CTCTCTTCAGCAGC	CAGCAGCAGCTGGGCGCGCCCTCTTGGCTGG	1089
765	AGGGGACCCGGCT	CCACACACACATCGCGCCCTTAGAGAG	CGCGCAGCCCATCTGGAGCA	824
1090	AGGGGACCCGGCT	CCACACACATCGCGCCCTTAGAGAG	CGCGCAGCCCATCTGGAGCA	1149
825	AAGAGAAAGGATAA	ACAGAAAGCACACCTCTCTAGGGT	TCCCAGGGGGCGCGGCTGGGG	884
1150	AAGAGAAAGGATAA	ACAGAAAGCACACCTCTCTAGGGT	TCCCAGGGGGCGCGGCTGGGG	1209
885	CTGGCTAGGTGAAA	AGGCAGAACACTCGCGCTCTCTTAGAGAG	GAGGAGTGTAGAGGAGGCG	944
1210	CTGGCTAGGTGAAA	AGGCAGAACACTCGCGCTCTCTTAGAGAG	GAGGAGTGTAGAGGAGGCG	1269
945	GGGGGCGCAGCA	CGCATCGTGTGGCCCTCCCTCC	CCACCTCCCTGTGTATAAATATTTA	1004
1270	GGGGGCGCAGCA	CGCATCGTGTGGCCCTCCCTCC	CCACCTCCCTGTGTATAAATATTTA	1329
1005	CATGTGATGTCT	GGTCTGAAATGCAACAGAGCT	TAAAGAGCTTGCAAAAAA	1057
1330	CATGTGATGTCT	GGTCTGAAATGCAACAGAGCT	TAAAGAGCTTGCAAAAAA	1382

RESULT 10

```

RESULTS 10
US-10-269-909-85
; Sequence 85, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-85

```

RESULT 11

RESUME 11
US-10-098-941-71
: Sequence 71, Application US/10098941
: Publication No. US2002019759A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Xu, Chongjun

193	Db	AAGAGTGCCTGTGGCCCTCGGAGAGCA CAGTGTCAAGGCA CGAAATCCAGAGCGCGAGG	252
321	Qy	TCTACGCCCGCCCTCGGCCCA CCGACCGCCTGGCCGTGCGCCCTTTCGCCCAAGCGGAGC	380
253	Db	TCTACGCCCGCCCTCGGCCCA CCGACCGCCTGGCCGTGCGCCCTTTCGCCCAAGCGGAGC	312
381	Qy	GCTTCCAGCGCTTCCAGCCGACCTATCCGTACCTGCGACGACGAGATCGACCTGCCGCCCA	440
313	Db	GCTTTCACACCGCTTCGAGCCCACTATCCGTACCTGCGACGACGAGATCGACCTGCCGCCCA	372
441	Qy	CCATCTCGCTGT CAGA CGGCGAGGAGCCGCCACCCCTA CCAGGCGCCCTTGCACTCTCCAGC	500
373	Db	CCATCTCGCTGT CAGA CGGCGAGGAGCCGCCACCCCTA CCAGGCGCCCTTGCACTCTCCAGC	432
501	Qy	TTGCGGACCCCGAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCAACA	560
433	Db	TTGCGGACCCCGAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCAACA	492
561	Qy	GAACCATCTTCGACAGTGCACCTGATGGATAGTGCACAGCTGGCGGGCCCTTGCCGCCCCCA	620
493	Db	GAACCATCTTCGACAGTGCACCTGATGGATAGTGCACAGCTGGCGGGCCCTTGCCGCCCCCA	552
621	Qy	GCAGTAACTCGGGCATCAGCGCCACGTGCTACCGGCGCGCGGCGGCATGGAGGGCGCGC	680
553	Db	GCAGTAACTCGGGCATCAGCGCCACGTGCTACCGGCGCGCGGCGGCATGGAGGGCGCGC	612
681	Qy	CGCCCACTACAGCGAGGTCAATCGGCCCACTACCCGGGGTCTCCTTCGACGACCAAGCAGA	740
613	Db	CGCCCACTACAGCGAGGTCAATCGGCCCACTACCCGGGGTCTCCTTCGACGACCAAGCAGA	672
741	Qy	GCAGTGGCGCGCCCTTCCTTGTGGAGGGGACCGCGCTCCACACACACACATCGCGGCCCC	800
673	Db	GCAGTGGCGCGCCCTTCCTTGTGTGGAGGGGACCGCGCTCCACACACACACATCGCGGCCCC	732
801	Qy	TAGAGCGCGACCATCTCGGAGCAAGAGAGAGTAAACAGAAAGGACACCTCTCTTAGG	860
733	Db	TAGAGCGCGACCATCTCGGAGCAAGAGAGAGTAAACAGAAAGGACACCTCTCTTAGG	792
861	Qy	GTCCCCAG	868
793	Db	GTCCCCAG	800

RESULT 13

```

US-09-934-249-1
:
: Sequence 1, Application US/0993249
: Patent No. US20020115081A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Lee, Richard T.
: APPLICANT: Landschulz, Katherine T.
: APPLICANT: Turi, Thomas G.
: APPLICANT: Thompson, John P.
: APPLICANT: Kennedy, Scott P.
:
: TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
:
: FILE REFERENCE: P0738/7001/ERR/KA
: CURRENT APPLICATION NUMBER: US/09/934,249
: CURRENT FILING DATE: 2001-08-21
: PRIOR APPLICATION NUMBER: US 60/227,159
: PRIOR FILING DATE: 2000-08-22
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1321
: TYPE: DNA
: ORGANISM: Homo Sapiens
:
: FEATURES:
:
: NAME/KEY: CDS
: LOCATION: (413)...(1273)
US-09-934-249-1

```

Query Match 70.6%; Score 749; DB 9; Length 1321;
Best Local Similarity 99.9%; Pred. No. 0;

RESULT 14

```

US-10-390-045-2
; Sequence 2, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKHEIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC AND
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995-0057-0000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18

```

Matches	799;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
y	105	CGAGCTCGAGTGTGTTTCAGATCATCATCATCGTGGTGTGATGATGTCATGTTGTGGTGG	164						
b	522	CGAGCTCGAGTGTGTTTCAGATCATCATCATCGTGGTGTGATGATGTCATGTTGTGGTGG	581						
y	165	TGATCACTGCTCTGAGCGACCTACAAAGCTGTCTGCAAGCTCTTTCATCAGACCGCGCAC	224						
b	582	TGATCACTGCTCTGAGCGACCTACAAAGCTGTCTGCAAGCTCTTTCATCAGACCGCGCAC	641						
y	225	GCAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAAGGATGCTGTGTGCCCTTCGAGGA	284						
b	642	GCAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAAGGATGCTGTGTGCCCTTCGAGGA	701						
y	285	GCACAGTGTCAAGCAACGGAATCCACAGACGCGAGGTCTACGCGCCGCCCTCGCGCCACCG	344						
b	702	GCACAGTGTCAAGCAACGGAATCCACAGACGCGAGGTCTACGCGCCGCCCTCGCGCCACCG	761						
y	345	ACCGCTGGCGCTGCGGCCCTTCGCCACAGCGGAGCGCTTCACACGCTTCACGCCACCT	404						
b	762	ACCGCTGGCGCTGCGGCCCTTCGCCACAGCGGAGCGCTTCACACGCTTCACGCCACCT	821						
y	405	ATCCGTACTCTCAGCAACAGATGCACTGCGCGCCACCATCTCTGCTGTCTAGAAGGGAGG	464						
b	822	ATCCGTACTCTCAGCAACAGATGCACTGCGCGCCACCATCTCTGCTGTCTAGAAGGGAGG	881						
y	465	AGCCCCCACTTACACAGGGCCCTGCACTCTCAGCTTTCGGGACCCCGAGCAGCAGCTGG	524						
b	882	AGCCCCCACTTACACAGGGCCCTGCACTCTCAGCTTTCGGGACCCCGAGCAGCAGCTGG	941						
y	525	AACTGAACCGGGAGTTCGCTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAACCTGA	584						
b	942	AACTGAACCGGGAGTTCGCTGCGCGCACCCCAACAGAACCATCTTCGACAGTGAACCTGA	1001						
y	585	TGATATGTGTCAGGCTGGGCGGCCCTTCGCCCTCCACGAGTAACTCTGGGATCAAGCGCCA	644						
b	1002	TGATATGTGTCAGGCTGGGCGGCCCTTCGCCCTCCACGAGTAACTCTGGGATCAAGCGCCA	1061						
y	645	CTGTCTACGGCAGCGGGCGCATGAGGGGCGCGCCACCTACAGCGAGGTCACTCG	704						
b	1062	CTGTCTACGGCAGCGGGCGCATGAGGGGCGCGCCACCTACAGCGAGGTCACTCG	1121						
y	705	GCACCTACCCGGGTCCTCTCTTCAGCAACAGACAGCAGTGGGCGGCCCTCTCTTGTGTGG	764						
b	1122	GCACCTACCCGGGTCCTCTCTCTTCAGCAACAGACAGCAGTGGGCGGCCCTCTCTTGTGTGG	1181						
y	765	AGGGACCCGGCTTCAACACACACATCGGCCCTCTAGAGAGCGCAGCATCTGAGGCA	824						
b	1182	AGGGACCCGGCTTCAACACACACATCGGCCCTCTAGAGAGCGCAGCATCTGAGGCA	1241						
y	825	AAGAGAAGGATAAACAAGAAAGAACACCTCTCTAGAGTTCCTCAGGGGGCGCGGCTGGGG	884						
b	1242	AAGAGAAGGATAAACAAGAAAGAACACCTCTCTAGAGTTCCTCAGGGGGCGCGGCTGGGG	1301						
y	885	CTGCGTAGGTGAAAAGGCAG	904						
b	1302	CTGCGTAGGTGAAAAGGCAG	1321						

4

Qy	761	CTGGAGGGACCCGGCTCCACACACATCGCGCCCTAGAGAGCGCGCCATCTGG	820
Db	661	CTGGAGGGGNCCTGGCTCCACACACATCGCGCCCTAGAGAGCGCGCCATCTGG	720
Qy	821	AGCAAGAGAGGATAAACAGAAAGGACCCCTCTCTAG	859
Db	721	AGCAAGAGAGGATAAACAGAAAGGACCCCTCTCTAG	759

Search completed: May 26, 2004, 22:14:31
Job time : 712 secs

POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY	576	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	635
DB	489	GTGACCTGATGATAGTCCAGGCTGGCGGCCCCCTTGCACCTTCCAGCTTGGGAGCCCCGAGC	430
QY	636	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	695
DB	429	TCAGCCGCCAGTCTACGGCAGCGGGGCGGACATGAGAGGGCGCGCCGCCACCTACAGCG	370
QY	696	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	755
DB	369	AGGTGATCGGCCACTACCGGGGTCTCTCTTCCAGCACACAGCAGCAGTGGGCGCGCTT	310
QY	756	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	815
DB	309	CCTTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	250
QY	816	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	875
DB	249	TCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTCTAGGGGTCCGCCAGGGGGCC	190
QY	876	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	935
DB	189	GGGCTGGGGCTGGCTAGGTGAAGAGGAGAACTCTCGCGCTTCTTAGAGAGGAGGTGAG	130
POLYA=Yes.		Location/Qualifiers	
1. .729		/organism="Homo sapiens"	
/mol_type="mRNA"		/db_xref="taxon:9606"	
/clone="UI-H-E21-bbg-h-14-0-UI"		/tissue_type="Chondrosarcoma Grade II"	
/dev_stage="Adult"		/lab_host="DH10B (Life Technologies)"	
/clone_lib="NCI CGAP Ch2"		/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
TAG TISSUE=grade-2-chondrosarcoma		TAG_LIB=UI-H-E21	
TAG_SEQ=ATCTAATATG"		68.0%; Score 722; DB 13; Length 729;	
Query Match		100.0%; Pred. No. 5.4e-172;	
Best Local Similarity		0; Mismatches 0; Indels 0; Gaps 0;	
Matches 722; Conservative			
QY	336	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	395
DB	729	GGCCACCGACCGCTTGGCGCTGCGCCCTTCCGCGGAGCGGCTTCCACCGCTTCC	670
QY	396	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	455
DB	669	AGCCACCTATCCGTACTCTGAGCAGCAGATCCACTGCGCCACCATCTCGCTGTCTAG	610
QY	456	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	515
DB	609	ACGGGAGGAGGCCCCACCTTACAGGGGCCCCCTTGACCTTCCAGCTTGGGAGCCCCGAGC	550
QY	516	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	575
DB	549	AGCAGCTGGAACCTGAACCGGAGTCTGGTGGCGGACCCGCCAACAGAACCATCTTCGACA	490
QY			

[illegible]

RESULT 3					
BQ691705					
LOCUS	BQ691705	655 bp	mRNA	linear	EST 15-JUL-2002
DEFINITION	AGENCOURT_8046876 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6208939				
	5' mRNA sequence.				

ACCESSION	BQ691705
VERSION	BQ691705.1
KEYWORDS	GI-21817021
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 655)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2367 row: g column: 20
High quality sequence stop: 645.
Location/Qualifiers
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208939"
FEATURES
source

```

/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/note="Organ: pâncreás; Vector: pOTe7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

ORIGIN	Query Match	61.5%;	Score 652;	DB 13;	Length 655;
	Best Local Similarity	100.0%;	Pred. No. 2.1e-154;		
	Matches 652;	Conservative	0;	Mismatches	0;
				Indels	Gaps
QY	410	TACCTGCAGCAGAGATCGACCTGCGCGCCACCATCTCGCTGTTCAGACGGGAGGAGCCC	469		
Db	1	TACCTGCAGCAGAGATCGACCTGCGCGCCACCATCTCGCTGTTCAGACGGGAGGAGCCC	60		
QY	470	CCACCTTACCAGGCGCCCTTCGACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGGAACCTG	529		
Db	61	CCACCTTACCAGGCGCCCTTCGACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGGAACCTG	120		
QY	530	AACCGGAGTTCGGTGGCGCGCACCCCAACAGAACATCTTCACACAGTCACTCGATCGGAT	589		
Db	121	AACCGGAGTTCGGTGGCGCACCCCAACAGAACATCTTCACACAGTCACTCGATCGGAT	180		
QY	590	AGTGCAGGCTGGCGCGCCCTTCGCCCCCGACGAGTAACTTCGGGCATTCAGCGCCACCTGTC	649		
Db	181	AGTGCAGGCTGGCGCGCCCTTCGCCCCCGACGAGTAACTTCGGGCATTCAGCGCCACCTGTC	240		
QY	650	TACGGCAGCGGGCGGCGCATGAGGGGCGCGCGCCACCTTACAGCGAGGTCATCGGCCAC	709		
Db	241	TACGGCAGCGGGCGGCGCATGAGGGGCGCGCGCCACCTTACAGCGAGGTCATCGGCCAC	300		
QY	710	TACCCGGGGTCTCTCTTCAGACACAGCAGAGCAGTGGGCGCGCTCTTGTCTGGAGGGG	769		
Db	301	TACCCGGGGTCTCTCTTCAGACACAGCAGAGCAGTGGGCGCGCTCTTGTCTGGAGGGG	360		
QY	770	ACC CGGCTTCACACACACATTCGGCGCCCTTAGAGAGCGCAGCCATCTTGGAGCAAGAG	829		
Db	361	ACC CGGCTTCACACACACATTCGGCGCCCTTAGAGAGCGCAGCCATCTTGGAGCAAGAG	420		
QY	830	AGGATTAACAGAAAGGACACCTCTCTAGGCTCCCGACGGGCGCGGCTGGGGCTGGC	889		
Db	421	AGGATTAACAGAAAGGACACCTCTCTAGGCTCCCGACGGGCGCGGCTGGGGCTGGC	480		
QY	890	TAGTGAAGAGGACAGAACTCCGCGCTCTTTAGAAAGAGAGTGAGAGGAAGCGCGGGGG	949		
Db	481	TAGTGAAGAGGACAGAACTCCGCGCTCTTTAGAAAGAGAGTGAGAGGAAGCGCGGGGG	540		
QY	950	CGCAGCAACGCATTCGTTGGCCCTTCCCTCCACCTCTCTGTGTATAAATAATTATCATGT	1009		
Db	541	CGCAGCAACGCATTCGTTGGCCCTTCCCTCCACCTCTCTGTGTATAAATAATTATCATGT	600		
QY	1010	GATGTCGTCTGAATGCAACAGCTTACAGAGCTTCGAAAAA	1061		
Db	601	GATGTCGTCTGAATGCAACAGCTTACAGAGCTTCGAAAAA	652		

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1. (bases 1 to 951)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12903836.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DK06NPL&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ015DK06NPL.
Location/Qualifiers
1. .951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YF12"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 61.5%; Score 652; DB 9; Length 951;
Best Local Similarity 99.6%; Pred. No. 1.6e-154;
Matches 802; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 168 TCACGCTGCTGTGAGCACTACAGCTGTCTGACGCTCTTCATCAGCGGCGACGACC 227
DB 846 TCACGCTGCTGTGAGCACTACAGCTGTCTGACGCTCTTCATCAGCGGCGACGACC 787

QY 228 AGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGCGCCCTCGGAGACA 287
DB 786 AGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGCGCCCTCGGAGACA 727

QY 288 CAGTGTACGAGCAACGAGATCCAGAGCCGAGGTCTACCGCGGCTTCGCGCCACCGACC 347
DB 726 CAGTGTACGAGCAACGAGATCCAGAGCCGAGGTCTACCGCGGCTTCGCGCCACCGACC 667

QY 348 GCCTGCGCTGCGCCCTTCGCGCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATC 407
DB 666 GCCTGCGCTGCGCCCTTCGCGCCAGCGGAGCGCTTCACCGCTTCAGCGCCACCTATC 607

QY 408 CTTACTCTCAGCAGCAGATCGACCTCTGCGCCCAACATCTGCTGTGACAGCGGAGGAGC 467
DB 606 CTTACTCTCAGCAGCAGATCGACCTCTGCGCCCAACATCTGCTGTGACAGCGGAGGAGC 547

QY 468 CCCCACCTTACAGAGGCGCCCTTGACCTCTGAGCTTCGGAGCCCGAGCGAGCTGGAC 527
DB 546 CCCCACCTTACAGAGGCGCCCTTGACCTCTGAGCTTCGGAGCCCGAGCGAGCTGGAC 487

QY 528 TGAACCGGAGTGTGCGCGCACCCCAACAGACACCATCTTCGACAGTGCCTCATCG 587
DB 486 TGAACCGGAGTGTGCGCGCACCCCAACAGACACCATCTTCGACAGTGCCTCATCG 427

QY 588 ATAGTGCCAGGCTGGCGGCGCCCTGCGCCCAACAGCAGTAACTCGGCGCATCGCCACT 647
DB 426 ATAGTGCCAGGCTGGCGGCGCCCTGCGCCCAACAGCAGTAACTCGGCGCATCGCCACT 367

QY 648 GCTACGCGAGCGCGCGGCGATGAGAGGGCGCGCGCCCTTACAGCGAGGTTCATCGGCC 707
DB 366 GCTACGCGAGCGCGCGGCGATGAGAGGGCGCGCGCCCTTACAGCGAGGTTCATCGGCC 307

QY 708 ACTACCGCGGCTCTCTTCCAGCACACAGCAGCAGTGGCGCGCCCTCTTGTGTGAGG 767
DB 306 ACTACCGCGGCTCTCTTCCAGCACACAGCAGCAGTGGCGCGCCCTCTTGTGTGAGG 247

QY 768 GGACCGCGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGGAGCAAG 827
DB 246 GGACCGCGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGGAGCAAG 187

QY 828 AGAAGGATAAAGAGAGAGAGACACCTCTCTAGGGTCCCGAGGGCGCGCGCTGGGGCTG 887
DB 186 AGAAGGATAAAGAGAGAGAGACACCTCTCTAGGGTCCCGAGGGCGCGCGCTGGGGCTG 127

QY 888 CTTAGTGTGAAAAGGAGAGACACTCCGCGCTCTTTAGAAGAGAGTGTAGAGCAAGCGCGG 947
DB 126 CTTAGTGTGAAAAGGAGAGACACTCCGCGCTCTTTAGAAGAGAGTGTAGAGCAAGCGCGG 67

QY 948 GGCGCAGCAACGATCGTGTGGCC 972
DB 66 GGCGCAGCAACGATCGTGTGGCC 42

RESULT 5

BQ636742
LOCUS
DEFINITION
BQ636742
609 bp mRNA linear EST 15-JUL-2002
Homo sapiens y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone hdi3h06 5', mRNA sequence.

ACCESSION

BQ636742.1 GI:21761201
VERSION
BQ636742
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE

1 (bases 1 to 609)
Wistow, G., Bernshtein, S.L., Wyatt, M.K., Ray, S., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), 196-204 (2002)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: b column: 06
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers

FEATURES

source

1. .609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hdi3h06"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMD108"
/clone_lib="Human Retina cDNA (Un-normalized,
unamplified): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I

primer-adaptor
[5'-pGACTAGTCTTAGATCGGAGCGCGCC(T)15-3']. EST analysis
was performed on the unamplified library at the NIH
Intramural Sequencing Center (NISC).

ORIGIN	Query Match	Best Local Similarity	Score	DB 13	Length	609;	Indels	0;	Gaps	0;
	Matches	609;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	214	CAGCGGCGACAGCGCGGCGGAGGAGAGAGATGCCCTCTCTCAGAGGATGCCCTGTG	273							
Db	1	CAGCGGCGACAGCGCGGCGGAGGAGAGATGCCCTCTCTCAGAGGATGCCCTGTG	60							
Qy	274	GCCCTCGGAGAGCAGTGTTCAGGCAACGAAATCCAGAGCGCGAGTCTACGCGCCGCC	333							
Db	61	GCCCTCGGAGAGCAGTGTTCAGGCAACGAAATCCAGAGCGCGAGTCTACGCGCCGCC	120							
Qy	334	TGCGCCACCGACCGCTGCGCGCTTGGCGCCCTTGGCGCCAGCGGAGCGCTTCCACCGCTT	393							
Db	121	TGCGCCACCGACCGCTGCGCGCTTGGCGCCCTTGGCGCCAGCGGAGCGCTTCCACCGCTT	180							
Qy	394	CGAGCCCACTATTCGCTACCTGCGACGACGAGATGCGCTCGCGCCACGATCTGCTGTC	453							
Db	181	CGAGCCCACTATTCGCTACCTGCGACGACGAGATGCGCTCGCGCCACGATCTGCTGTC	240							
Qy	454	AGACGGGAGGAGCGCCACCTTACAGAGGCGCCCTGCAACCTTCCAGCTTCGGGACCCGA	513							
Db	241	AGACGGGAGGAGCGCCACCTTACAGAGGCGCCCTGCAACCTTCCAGCTTCGGGACCCGA	300							
Qy	514	GCAGCAGCTGGAACTGAACCGGAGTGGTGGCGGCAACCCCAACAGAACCACTTTGGA	573							
Db	301	GCAGCAGCTGGAACTGAACCGGAGTGGTGGCGGCAACCCCAACAGAACCACTTTGGA	360							
Qy	574	CAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	633							
Db	361	CAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	420							
Qy	634	CATCAGCGCCACGCTGTACGCGAGCGCGCGCGCATGAGGGGCGCGCGCCACCTACAG	693							
Db	421	CATCAGCGCCACGCTGTACGCGAGCGCGCGCGCATGAGGGGCGCGCGCCACCTACAG	480							
Qy	694	CGAGGTGATCGGCACTACCGCGGCTCTCTCTTCCAGGACCGAGCAGAGTGGCGCGCC	753							
Db	481	CGAGGTGATCGGCACTACCGCGGCTCTCTCTTCCAGGACCGAGCAGAGTGGCGCGCC	540							
Qy	754	CTCCTTGTGAGGGGACCGGCTTCCACACACACACACACACACACACACACACACAC	813							
Db	541	CTCCTTGTGAGGGGACCGGCTTCCACACACACACACACACACACACACACACACAC	600							
Qy	814	CATCTGGAG 822								
Db	601	CATCTGGAG 609								

RESULT 6
LOCUS BQ690750
DEFINITION AGENCOURT_8046394 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209341
5', mRNA sequence.
ACCESSION BQ690750
VERSION BQ690750.1 GI:21816066
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LMC2368 row: h column: 14
High quality sequence stop: 627.
Location/Qualifiers
1. .890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="INAGR:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GSCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN	Query Match	Best Local Similarity	Score	DB 13	Length	890;	Indels	0;	Gaps	0;
	Matches	651;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	TCCTCTCTGGTTCGGTGAAAGCGCTTGGGGGTTTCACTGGGCGCATGATCCCGAGCTGC	60							
Db	24	TCCTCTCTGGTTCGGTGAAAGCGCTTGGGGGTTTCACTGGGCGCATGATCCCGAGCTGC	83							
Qy	61	TGAGAACTGAAGCGGAGCGGCTCTCTTCGGAACACAGGCAATGCGGAGCTGGAGTTGT	120							
Db	84	TGAGAACTGAAGCGGAGCGGCTCTCTTCGGAACACAGGCAATGCGGAGCTGGAGTTGT	143							
Qy	121	TCAGATCATCATCATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	180							
Db	144	TCAGATCATCATCATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	203							
Qy	181	GAGCCACTACAAGCTGTCTGACGCTCTTTCATCAGCGCGCACAGCCAGCGCGGAGAG	240							
Db	204	GAGCCACTACAAGCTGTCTGACGCTCTTTCATCAGCGCGCACAGCCAGCGCGGAGAG	263							
Qy	241	AGAAGATGCGCTGTCTCTCAGAGGATGCTCTGTGGCCCTCGGAGAGCAAGTGTGAGCAA	300							
Db	264	AGAAGATGCGCTGTCTCTCAGAGGATGCTCTGTGGCCCTCGGAGAGCAAGTGTGAGCAA	323							
Qy	301	CGGAATCCAGAGCGCGAGTCTACCGCGCGCTCGGCGCACCGAGCGCTGGCGGTGCC	360							
Db	324	CGGAATCCAGAGCGCGAGTCTACCGCGCGCTCGGCGCACCGAGCGCTGGCGGTGCC	383							
Qy	361	GCCTTTGGCGCGAGCGGCGCTTTCACCGCTTTCAGCGCCACCTATCCGTAACCTGACGA	420							
Db	384	GCCTTTGGCGCGAGCGGCGCTTTCACCGCTTTCAGCGCCACCTATCCGTAACCTGACGA	443							
Qy	421	CGAGATGACCTGCGCGCCACCATCTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	480							
Db	444	CGAGATGACCTGCGCGCCACCATCTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	503							
Qy	481	GGGCGGCTGACCGCTTCAGCTTGGGACCGCGAGCAGCTGGACTGAACTGAACTGGAGTGC	540							
Db	504	GGGCGGCTGACCGCTTCAGCTTGGGACCGCGAGCAGCTGGACTGAACTGAACTGGAGTGC	563							
Qy	541	GGTGGCGGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	600							
Db	564	GGTGGCGGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	623							
Qy	601	GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	652							

pool of mRNA from 2 cell lines from Enchondroma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGGTCACTC. The cell lines were provided by Dr. James Martin from the University of Iowa. TAG TISSUE=Enchondroma cell line (Mix of EN1 and EN2) TAG LIB=UI-H-FGI TAG_SEQ=CGGTCACTC"

ORIGIN

Query Match 55.1%; Score 585; DB 13; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.5e-137;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 477 ACCAGGGCCCTGACCCCTCAGCTTCGGGACCCGAGCAGCAGCTGGAACTGAACCGG 536
DB 588 ACCAGGGCCCTGACCCCTCAGCTTCGGGACCCGAGCAGCAGCTGGAACTGAACCGG 529

QY 537 AGTCGGTGGCGGACACCCCAACAGAACCACTTTCGACAGTCACTGATGATAGTGCCA 596
DB 528 AGTCGGTGGCGGACACCCCAACAGAACCACTTTCGACAGTCACTGATGATAGTGCCA 469

QY 597 GGCTGGGCGGCCCTGCGGCCCGGACAGTAACCTGGGCAATCAGCGCCACCTGCTACGGCA 656
DB 468 GGCTGGGCGGCCCTGCGGCCCGGACAGTAACCTGGGCAATCAGCGCCACCTGCTACGGCA 409

QY 657 GCGGCGGGCGCATGGAGGGCGCGGCCCACTTACAGAGGTCATCGGCCCACTACCGCG 716
DB 408 GCGGCGGGCGCATGGAGGGCGCGGCCCACTTACAGAGGTCATCGGCCCACTACCGCG 349

QY 717 GGTCTCTCTTCCAGCACCAGCAGCAGTGGCGCGCTCTTCTTGGAGGGACCGCGC 776
DB 348 GGTCTCTCTTCCAGCACCAGCAGCAGTGGCGCGCTCTTCTTGGAGGGACCGCGC 289

QY 777 TCACACACACACATCGCGCCCTTAGAGAGCGCAGCCTCTGGAGCAAGAAGAGATA 836
DB 288 TCACACACACACATCGCGCCCTTAGAGAGCGCAGCCTCTGGAGCAAGAAGAGATA 229

QY 837 AACAGAAAGGACACCTCTAGGTGTCCTCCAGGGGGCGCGGCTGGGCTGCGTAGGTGA 896
DB 228 AACAGAAAGGACACCTCTAGGTGTCCTCCAGGGGGCGCGGCTGGGCTGCGTAGGTGA 169

QY 897 AAAGGCGAGAACACTCGCGCGCTCTTAGAGAGGAGTGAGAGGAAGCGGGGGCGCAGCA 956
DB 168 AAAGGCGAGAACACTCGCGCGCTCTTAGAGAGGAGTGAGAGGAAGCGGGGGCGCAGCA 109

QY 957 AGCATGCTGTGGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCT 1016
DB 108 AGCATGCTGTGGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCT 49

QY 1017 GGTCTGAATGCAACAGTAAAGAGCTTGCARAAAAA 1061
DB 48 GGTCTGAATGCAACAGTAAAGAGCTTGCARAAAAA 4

RESULT 9

BU730650/c
LOCUS BU730650 629 bp mRNA linear EST 09-OCT-2002
DEFINITION UI-E-C11-afq-o-09-0-UI.s1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION BU730650
VERSION BU730650.1 GI:23654753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 629)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source

1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afq-o-09-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: p7773-Pac (Pharmacia)-with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=RPE and Choroid
TAG LIB=UI-E-C11
TAG_SEQ=ACCTA"

ORIGIN

Query Match 53.7%; Score 570; DB 13; Length 629;
Best Local Similarity 100.0%; Pred. No. 8.2e-134;
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 488 TGCACCCCTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTGGTGGCG 547
DB 577 TGCACCCCTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTGGTGGCG 518

QY 548 GCACCCCAACAGAACCACTTTCGACAGTCACTGATGATAGTGCAGGCTGGGGCGG 607
DB 517 GCACCCCAACAGAACCACTTTCGACAGTCACTGATGATAGTGCAGGCTGGGGCGG 458

QY 608 CCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACCGTGTCTACGCGAGCGGGCGGCG 667
DB 457 CCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACCGTGTCTACGCGAGCGGGCGGCG 398

QY 668 ATGGAGGGGGCGCGGCCACCTTACAGGAGGTATGCGGCCCATCTACCCGGGGTCTCTTTC 727
DB 397 ATGGAGGGGGCGCGGCCACCTTACAGGAGGTATGCGGCCCATCTACCCGGGGTCTCTTTC 338

Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

337	CAGCAACGACGAGCATGTGGGCGCCCTCTTCTGGAGGGGACCCGGCTCCACCAACA	278
788	CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGA	847
277	CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGA	218
848	GACCCCTCTCTAGGGTCCCAAGGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGAGGCGAGAAC	907
217	CACCCCTCTCTAGGGTCCCAAGGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGAGGCGAGAAC	158
908	ACTCCGCGCTTCTTAGAGAGGAGTGTAGAGAGGAGCGCGGGGGCGCAGCAACCGCATCGTGT	967
157	ACTCCGCGCTTCTTAGAGAGGAGTGTAGAGAGGAGCGCGGGGGCGCAGCAACCGCATCGTGT	98
968	GGGCCCTCCCTCCCACTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGC	1027
97	GGCCCTCCCTCCCACTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGC	38
1028	ACAAGCTAAGAGAGCTTGCAAAAAA	1057
37	ACAAGCTAAGAGAGCTTGCAAAAAA	8

RESULT 10	LOCUS	DEFINITION	559 bp	mRNA	linear	EST 28-SEP-2000
BB855409/c	7913f05.x1	NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3306369 3' similar to TR:Q9UJD3 Q9UJD3 DJ718J7.1 ; mRNA sequence.				

ACCESION	BE855409				
VERSION	BE855409.1	GI:10367404			
KEYWORDS	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 559)				
AUTHORS	NCI/NINDS-CGAP Institute (http://www.ncbi.nlm.nih.gov/ncicgap).				
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/ETGAP), Tumor Gene Index				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Robert Strausberg, Ph. D. Email: CGAPbs-remail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph. D.				

RESULT 11
BM974296/C

LOCUS	Accession	Size	Source	EST
LOCUS	BM974296	626 bp	mRNA	EST 20-FEB-2003
DEFINITION	UI-CP-EC1-aca-k-23-0-UI.s1 UI-CP-EC1 Homo sapiens cDNA clone			
DESCRIPTION	UI-CP-EC1-aca-k-23-0-UI 3', mRNA sequence.			

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

ORGANISM HOMO SAPIENS

REFERENCE
AUTHORS
TITLE
JOURNAL

Query Match 52.7%; Score 559; DB 10; Length 559;
Best Local Similarity 100.0%; Pred. No. 5.2e-131;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Caps 0;

488	TGCA	CCCTCCAGCTTTCGGGA	CCCCGAGCAGCAGCTGGAACTGAA	CCGGGAGTCCGGTGGCGC	547
559	TGC	ACCTTCAGCTTTCGGGA	CCCCGAGCAGCAGCTGGAACTGAA	CCGGGAGTCCGGTGGCGC	500
548	GC	ACCCCAAACAGAACCATCTTTCGACAGTGA	CCTGATGGATAGTGC	CAGGCTGGCGCGC	607
459	GC	ACCCCAAACAGAACCATCTTTCGACAGTGA	CCTGATGGATAGTGC	CAGGCTGGCGCGC	440
608	CCCTG	CCCCCAGCAGTAACTCTGGGCGATC	AGCGCCACGTGCTAACGGCAGCGCGGCGC	667	
439	CCCTG	CCCCCAGCAGTAACTCTGGGCGATC	AGCGCCACGTGCTAACGGCAGCGCGGCGC	380	
668	ATG	GAGGGCGCGCGCCCACTTACAGCGAGGTGC	TGGCCACTTACCCGGGGTCTCTCTTC	727	
379	ATG	GAGGGCGCGCGCCCACTTACAGCGAGGTGC	TGGCCACTTACCCGGGGTCTCTCTTC	320	
728	GAGCA	CCAGCAGAGCAGTGGCGCGCTCTTGTCTG	TGGAGGGGAGCCCGGCTCCACCA	CACA	787
319	CAGCA	CCAGCAGCAGTGGCGCGCTCTTGTCTG	TGGAGGGGAGCCCGGCTCCACCA	CACA	260
788	CACAT	CGCGCCCTTAGAGAGCGCAGCCATCTCG	GACAAAGAGAGATAAACAGAAAGGA	847	
259	CACAT	CGCGCCCTTAGAGAGCGCAGCCATCTCG	GACAAAGAGAGATAAACAGAAAGGA	200	
848	CAC	CTCTCTAGGGTCCCCAGGGGGCGGGCTG	CGGGCTGGCGGTGCAAAAGCGAGAAC	907	
159	CAC	CTCTCTAGGGTCCCCAGGGGGCGGGCTG	CGGGCTGGCGGTGCAAAAGCGAGAAC	140	
908	ACT	CCGCGCTCTTAGAGAGAGGTGAGAGGAG	CGGGGGCGGCGAGCAACGATCTGT	967	
139	ACT	CCGCGCTCTTAGAGAGAGGTGAGAGGAG	CGGGGGGGCGGCGAGCAACGATCTGT	80	
968	G	CCCTCCCTCCCACTCCCTGTGTATAAATAT	TTACATGTGATCTCTGGTCTGAATGC	1027	
79	G	CCCTCCCTCCCACTCCCTGTGTATAAATAT	TTACATGTGATCTCTGGTCTGAATGC	20	
1028	ACA	GCTAAGAGAGCTTGC	1046		
19	ACA	GCTAAGAGAGCTTGC	1		

RESULT 11
BM974296/C

LOCUS	Accession	Size	Source	EST
LOCUS	BM974296	626 bp	mRNA	EST 20-FEB-2003
DEFINITION	UI-CP-EC1-aca-k-23-0-UI.s1 UI-CP-EC1 Homo sapiens cDNA clone			
FEATURES	UI-CP-EC1-aca-k-23-0-UI 3', mRNA sequence.			

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

ORGANISM HOMO SAPIENS

REFERENCE
AUTHORS
TITLE
JOURNAL

QY 926 GAGGAGTGGAGGAAGCGCGGGGGCCGCAGCAACGCCATCGTGTCGTGGCCCTCCCCCTCCACCT 985
|||
Db 139 GAGGAGTGGAGGAAGCGCGGGGGCCGCAGCAACGCCATCGTGTCGTGGCCCTCCCCCTCCACCT 80
|||
QY 986 CCTGTCTATAAATATTTCATGTCGATGTCGTGGTCTGAATGCACAAGCTTAAGAGAGCTTG 104
|||
Db 79 CCTGTCTATAAATATTTCATGTCGATGTCGTGGTCTGAATGCACAAGCTTAAGAGAGCTTG 20
|||
QY 1046 CAAAAAAAAAAAAAA 1061
|||||
Db 19 CAAAAAAAAAAAAAA 4
|||||

RESULT 12	BQ575582/c
LOCUS	
DEFINITION	570 bp mRNA linear EST 19-JUN-2002 UT-H-E21-bbf-f-09-0-UI.s1 NCI CGAP_Ch2 Homo sapiens cDNA clone UT-H-E21-bbf-f-09-0-UI 3', mRNA sequence.
ACCESSION	BQ575582
VERSION	BQ575582.1 GI:21478899
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 570)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of Orthopedics cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu Seq primer: M13 FORWARD POLYA=Yes.

```

FEATURES
source
    Location/Qualifiers
    1..570
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="UI-H-E21-bbf-f-09-0-UI"
        /tissue_type="Chondrosarcoma Grade II"
        /dev_stage="Adult"
        /lab_host="PHI08 (Life Technologies)"
        /clone_lib="NCI CGAP Ch2"
        /notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."
        TAG_TISSUE=grade-2-Chondrosarcoma
        TAG_LIB=UI-H-E21
        TAG_SEQ=ATCTAATATG"

```

ORIGIN

Query Match 50.5%; Score 536; DB 13; Length 570;
 Best Local Similarity 100.0%; Pred. NO. 3e-125;

5', mRNA sequence.
 CB049800
 VERSION CB049800.1 GI:27788087
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 563)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbe-remail.nih.gov
 cDNA Library Preparation:
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM8009 row: E column: 1
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1. 563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3271656"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Pr28"
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 110192-110199, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES

source
 1. 563
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3271656"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Pr28"
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 110192-110199, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 50.0%; Score 530; DB 14; Length 563;
 Best Local Similarity 100.0%; Pred. No. 9.7e-124; Indels 0; Gaps 0;
 Matches 530; Conservative 0; Mismatches 0;
 528 TGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 587
 10 TGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 69
 588 ATAGTGCAGGCTGGGGGGCCCTGCCCCCAGCAGTAACTCGGGCATAGCCGACGT 647
 70 ATAGTGCAGGCTGGGGGGCCCTGCCCCCAGCAGTAACTCGGGCATAGCCGACGT 129
 648 GCTACGGCAGCGGGGGCGATGAGGGCGCGCCGACCTACAGCAGGTCTATCGGCC 707
 130 GCTACGGCAGCGGGGGCGATGAGGGGGCGCGCCGACCTACAGCAGGTCTATCGGCC 189
 708 ACTACCGGGGGTCTCTCTTCAGCACCAGCAGCAGTGGGGCCGCTCTCTCTGCTGGAGG 767
 190 ACTACCGGGGGTCTCTCTTCAGCACCAGCAGCAGTGGGGCCGCTCTCTCTGCTGGAGG 249
 768 GGAACCGGGCTCCACCAACACATCGGGCCCTTAGAGAGGCGAGCCATCTGAGGAGAAAG 827
 250 GGACCGGGCTCCACCAACACATCGGGCCCTTAGAGAGGCGAGCCATCTGAGGAGAAAG 309
 828 AGAAGGATAACAGAGGACACCTCTCTAGGGTCCCCCAGGGGGGGGGGGCTGGGGCTG 887
 310 AGAAGGATAACAGAGGACACCTCTCTAGGGTCCCCCAGGGGGGGGGGGCTGGGGCTG 369

QY 888 CGTAGGTGAAAAGGCGAGCACTCCGCGCTTCTTTAGAAAGAGAGTGAGAGGAGCGGGG 947
 DB 370 CGTAGGTGAAAAGGCGAGCACTCCGCGCTTCTTTAGAAAGAGAGTGAGAGGAGCGGGG 429
 QY 948 GGCGCAGCAAGCGCATCGTGGCCCTCCCTCCACCTCCCTCCCTGCTATATAATTTACAT 1007
 DB 430 GGCGCAGCAAGCGCATCGTGGCCCTCCCTCCACCTCCCTCCCTGCTATATAATTTACAT 489
 QY 1008 GTGATGCTGCTGCTGAATGCACAACTAAGAGAGCTTGCCAAAAA 1057
 DB 490 GTGATGCTGCTGCTGAATGCACAACTAAGAGAGCTTGCCAAAAA 539
 RESULT 15
 BO686793 844 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT_8345390 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248044
 DEFINITION 5', mRNA sequence.
 ACCESSION BO686793
 VERSION BO686793.1 GI:21812109
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 844)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM2387 row: e column: 05
 High quality sequence stop: 593.
 Location/Qualifiers
 1. 844
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6248044"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_110"
 /note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

Query Match 49.8%; Score 528; DB 13; Length 844;
 Best Local Similarity 99.8%; Pred. No. 2.3e-123;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 483 GCCCTTCACCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAAACCGGAGTGG 542
 DB 1 GCCCTTCACCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAAACCGGAGTGG 60
 543 TGCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGG 602
 DB 61 TGCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGG 120
 603 GGCGGCCCTGCCCCCAGCAGTAACTCGGGCATCAGGCCACCTGCTACGCGAGCGGG 662

```

Db      121  GCGGCCCTCGCCCGCCAGCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGCG 180
Qy      663  GCGGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCCT 722
Db      181  GCGGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCCT 240
Qy      723  CCTTCCAGCACACGACGAGCAGTGGGCGGCCCTCTCTGCTGGAGGGGACCCGGCTCCACC 782
Db      241  CCTTCCAGCACACGACGAGCAGTGGGCGGCCCTCTCTGCTGGAGGGGACCCGGCTCCACC 300
Qy      783  ACACACATCGCGCCCTTAGAGAGCGCGCATCTGGAGCAAAGAGAGATAAACAGA 842
Db      301  ACACACATCGCGCCCTTAGAGAGCGCGCATCTGGAGCAAAGAGAGATAAACAGA 360
Qy      843  AAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGCTAGGTGAAGGC 902
Db      361  AAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGCTAGGTGAAGGC 420
Qy      903  AGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGGCGGGGGCGCGCAACGCAT 962
Db      421  AGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGGCGGGGGCGCGCAACGCAT 480
Qy      963  CGTGTGGCCCTCCCGCTCCCGCTCCCTGTGTATATAATTTACATGTGTCTGTGTCTG 1022
Db      481  CGTGTGGCCCTCCCGCTCCCGCTCCCTGTGTATATAATTTACATGTGTCTGTGTCTG 540
Qy      1023  AATGCACAGCTAAGAGAGCTTGCACAAAAA 1061
Db      541  AATGCACAGCTAAGAGAGCTTGCACAAAAA 579

```

Search completed: May 26, 2004, 20:25:25
 Job time : 4421 secs